Analysis of Sparse and Noisy Ocean Current Data Using Flow Decomposition. Part II: Applications to Eulerian and Lagrangian Data

PETER C. CHU AND LEONID M. IVANOV

Naval Postgraduate School, Monterey, California

TATIANA P. KORZHOVA, TATIANA M. MARGOLINA, AND OLEG V. MELNICHENKO

*Marine Hydrophysical Institute and Center for Ecological Modeling and Recreation Geography,
National Ukrainian Academy of Sciences, Sevastopol, Crimea, Ukraine*

ABSTRACT

The capability of the reconstruction scheme developed in Part I is demonstrated here through three practical applications. First, the nonlinear regression model is used to reproduce the upper-layer three-dimensional circulation of the eastern Black Sea from model data distorted by white and red noises. Second, the quasigeostrophic approximation is used to reconstruct the shallow water circulation pattern in an open domain with various sampling strategies. Third, the large-scale circulation in the Southern Ocean is reproduced from the First Global Atmospheric Research Program (GARP) Global Experiment (FGGE) drifter data with noncontrollable noise statistics. All three cases confirm that the theoretical approach is robust to various noise-to-signal ratios, number of observations, and station disposition. Using the simplified open boundary condition for analyzing long-term observational data is recommended because the nonlinear regression procedure requires considerable computer resources.

1. Introduction

In Part I (Chu et al. 2003), a theoretical approach was developed for reconstructing a three-dimensional velocity field from sparse and noisy data. The scheme can be used for reconstructing velocity field from data in Eulerian and Lagrangian frames. To demonstrate the usefulness of this approach and to test the sensitivity of the approach to possible variations of noise-to-signal ratio, station disposition, and observational sample size, the schemes are applied to reconstruct the circulation in the eastern Black Sea, a rectangular open domain, and the Southern Ocean. The first two cases (identical twin experiments) represent the reconstruction of velocity from Eulerian "observations," and the third one (using observational data) represents the reconstruction of velocity from Lagrangian data.

The first experiment is designed to verify the model sensitivity to the selection of optimal model parameters (K_{opt}, M_{opt}, S_{opt}) through reconstructing the circulation in the eastern Black Sea. It is of considerable interest because the pipeline "Blue Flow" between Russia and Turkey will be built in the near future. Assessing the

ecological safety of the project has become urgent, and a realistic three-dimensional flow field is needed. Reconstruction of the three-dimensional baroclinic circulation in the eastern part of the Black Sea was performed with realistic topography and controlled noise in the observations. The spatial structure of the circulation was obtained using both scalar potentials Ψ and Φ.

The second experiment is designed to verify the model sensitivity to the sampling strategy through reconstructing circulation in a rectangular open domain with a flat bottom. Such an experiment has a practical significance. Presently, high-frequency radars are widely used in measuring circulation in the coastal ocean. Usually, the radar data are collected in a domain of limited size with high-resolution grids, which is very attractive if the circulation can be reconstructed in a domain larger than the radar polygon (Shulman et al. 2000). The question arises, how large should the domain be for a realistic flow reconstruction? In the experiment, we controlled the noise level in the data and used a simplified condition at the open boundary. Preliminary estimations demonstrated that, with given wind forcing, subgrid parameterization, and basin geometry, the shallow water circulation can be described using the quasigeostrophic streamfunction, which coincides with the toroidal potential Ψ.

The third experiment is designed to verify the scheme

Corresponding author address: Dr. Peter C. Chu, Naval Ocean Analysis and Prediction Lab, Department of Oceanography, Naval Postgraduate School, Monterey, CA 93943.
E-mail: chu@nps.nave.mil

	Report Documentation Page		*Form Approved* *OMB No. 0704-0188*

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **2003**	2. REPORT TYPE	3. DATES COVERED **00-00-2003 to 00-00-2003**

4. TITLE AND SUBTITLE **Analysis of Sparse and Noisy Ocean Current Data Using Flow Decomposition. Part II: Applications to Eulerian and Lagrangian Data**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Naval Postgraduate School,833 Dyer Road,Monterey,CA,93943**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
See Report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT **Same as Report (SAR)**	18. NUMBER OF PAGES **21**	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**			

Standard Form 298 (Rev. 8-98)
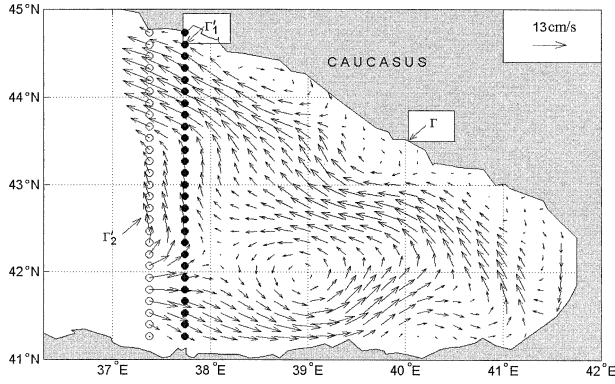Prescribed by ANSI Std Z39-18

FIG. 1. The exact velocity field at 20 m deep in the eastern Black Sea. Note that Γ_1' and Γ_2' are the open boundary and extended open boundary, respectively.
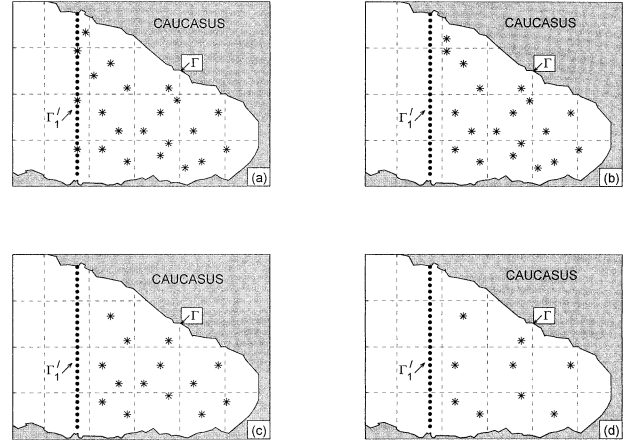

FIG. 2. Four station dispositions: (a) 22 stations with data at the open boundary, (b) 19 stations without data at the open boundary, (c) 14 stations without data at the open boundary, and (d) 9 stations without data at the open boundary.

using observational buoy data with unknown noise statistics and to examine the validity of boundary extension method through reconstructing the large-scale circulation in the Southern Ocean. More than 200 drifting buoys were deployed in the Southern Ocean during the First Global Atmospheric Research Program (GARP) Global Experiment (FGGE) from December 1978 to December 1979. Such a dataset provides useful information on the seasonal variability of circulation from 70° to 20°S (Hoffman 1985; Patterson 1985; Daniault and Menard 1985; Eremeev and Ivanov 1987). Here, we use the FGGE buoy data to reconstruct mean surface circulation in the Southern Ocean from 70° to 48°S on the 2° × 2° grids in May 1979. The oceanographic interpretation of the reconstructed seasonal variability of the circulation will be presented in another paper.

2. Case 1: Eastern Black Sea circulation

The first experiment is designed to determine the model sensitivity to the selection of optimal model parameters (K_{opt}, M_{opt}, S_{opt}) and to verify the boundary extension method through reconstructing the circulation in the eastern Black Sea.

a. Velocity "data"

A z-coordinate primitive model is used for the whole Black Sea with 14.6-km horizontal resolution and 21 vertical levels (Knysh et al. 2001) to obtain the three-dimensional circulation. The horizontal and vertical viscosities are taken as 10^7 cm^2 s^{-1} and 5 cm^2 s^{-1} for the entire water column. The horizontal diffusion coefficient is set to be 5×10^5 cm^2 s^{-1}, and the vertical diffusion coefficient varies with depth in the following way: 2.5 cm^2 s^{-1} above 5 m deep, 0.1 cm^2 s^{-1} between 5 and 15 m deep, and 0.03 cm^2 s^{-1} below 15 m deep. The model is enclosed by rigid lateral and bottom boundaries with no-slip and no heat and salt exchange conditions. The model is integrated for 25 yr from the rest with monthly mean surface fluxes of momentum, heat, and

water mass and with assimilation of climatological monthly mean three-dimensional temperature and salinity data (Eremev et al. 1992; Staneva et al. 1995); see Knysh et al. (2001) for detailed information about the model and numerical integration.

The final solutions show a quasiperiodic characteristic in the circulation pattern. We use the three-dimensional circulation calculated at 20 m deep on 15 December as the exact "observational" velocity field, which shows a cyclonic gyre centered near 42°N, 39°E in the eastern Black Sea with a jet stream (rim current) and several small anticyclonic eddies (around 100 km in diameter) around it (Fig. 1). This simulated "observational dataset," treated as "perfect" (no error), is used as a data source.

b. Controlled noise

Four different sampling strategies A, B, C, and D (Fig. 2) are used to obtain data from a velocity data source (Fig. 1). Notice that there are no "observations" at the open boundary Γ_1' for B–D data sampling. The station dispositions are designed to investigate the effect of sampling distribution and density on the reconstruction quality.

An error (noise) is added to the sampled observational data. The noise is represented by a two-dimensional, isotropic, incompressible, and homogeneous Gaussian-type field with the correlation function

$$G(\mathbf{r} - \mathbf{r}') = I^2 \exp\left[-\frac{(\mathbf{r} - \mathbf{r}')^2}{L_{cor}^2}\right]. \qquad (2.1)$$

Here, $\mathbf{r} = x\mathbf{i} + y\mathbf{j}$ is the displacement vector; L_{cor} and I are the correlation radius and intensity of noise, respectively; $L_{cor} \to 0$ for the white noise and otherwise for the red noise. The stochastic field with the corre-
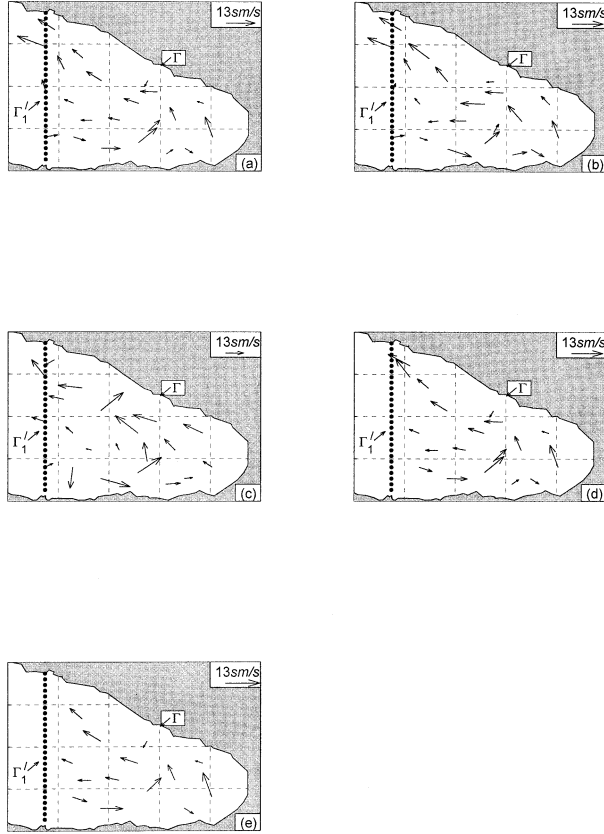
FIG. 3. Typical observational data collected from (a) disposition A with a white noise ($\eta_1 = 0.3$), (b) disposition A with a color noise ($\eta_1 = 0.3$, $L_{cor} = 45$ km), (c) disposition A with a strong white noise ($\eta_1 = 1.2$), (d) disposition B with a white noise ($\eta_1 = 0.3$), and (e) disposition C with a white noise ($\eta_1 = 0.3$).

lation function (2.1) was used in Sabelfeld's (1991) technique. The observations are distorted by various noises.

Three typical sets of sampled data (A, B, C) with white and red noises are considered (Fig. 3). Notice that if high-intensity noise is introduced into the perfect observations, the reconstruction process becomes weakly sensitive to whether there are observations along the open boundary or not.

c. Noise-to-signal ratio

Two types of error affect the noise-to-signal ratio η_1: stochastic noise in the distorted observations, and mode truncation [see (3.3) in Part I]. The second type of error—that is, cutoff of high-order modes—is regarded as the deterministic perturbation.

The noise-to-signal ratio is represented by (Eykhoff 1973):

$$\eta_1 = \frac{I + \sqrt{E_{cut}}}{\sqrt{E}}, \qquad (2.2)$$

where E and E_{cut} are the total and truncated horizontal circulation energies, respectively. In the numerical ex-

periments η_1 varies from 0.1 to 1.2. When the toroidal (poloidal) potential is truncated at around 20 (3), the contribution of the deterministic perturbation to η_1 may be neglected.

d. Toroidal and poloidal components

In the reconstruction, the open boundary extension method was used (Fig. 1). With boundary conditions given along Γ_2', 200 basis functions are calculated for both scalar potentials using MATLAB (1997). The numerical procedure for constructing the rotation matrix **R** is discussed by Ivanov et al. (2001). The integral equation (3.5) of Part I is reduced to linear algebraic equations and solved using a package of computational programs from Tikhonov et al (1990). Here, this approach is compared with the smoothed splines only.

For scalar ocean fields, the regularization technique in a fitting model is compared with the optimal interpolation (OI), the spline interpolation based on cross validation (CV), generalized cross validation (GCV), and Tikhonov's approach (Ivanov et al. 2001). Shultz (1999) compared the reconstruction technique based on the mode decomposition (Eremeev et al. 1992) with OI for reconstructing surface currents on the Louisiana–Texas shelf.

e. Optimal truncation

A question that needs to be answered before reconstruction is how to select the optimal model parameters: K_{opt} and M_{opt}, the number of toroidal and poloidal modes; and S_{opt}, the order of continuous polynomial used to parameterize the scalar function $\kappa(\tau)$. Vapnik–Chervonkis cost function (VCCF) is used to determine $(K_{opt}, M_{opt}, S_{opt})$ as explained in section 5a of Part I. For disposition A and B, the ranges of truncation parameters are found:

$$K_{opt} = 15 - 23, \qquad M_{opt} = 1 - 4, \quad \text{and} \quad S_{opt} = 5,$$

which does not vary with the data noise level noticeably. For simplicity, the selection of K_{opt} and M_{opt} is demonstrated using the empirical (J_{emp}), VCCF (J_{NM}), and statistical $\langle J \rangle$ cost functions for a given $S_{opt} = 5$, with the observational data distorted by 30% white noise (Fig. 3a).

1) EMPIRICAL COST FUNCTION

Figure 4a shows the dependence of the empirical functional J_{emp} on the mode truncation parameters K and M. A maximum value of J_{emp} is found at $(K_{opt}, M_{opt}) = (19, 4)$. Using

$$(K_{opt}, M_{opt}, S_{opt}) = (19, 4, 5),$$

the reconstructed field has an unrealistic jet near Caucasus, which indicates the overdetermination of the poloidal component in the general circulation pattern.
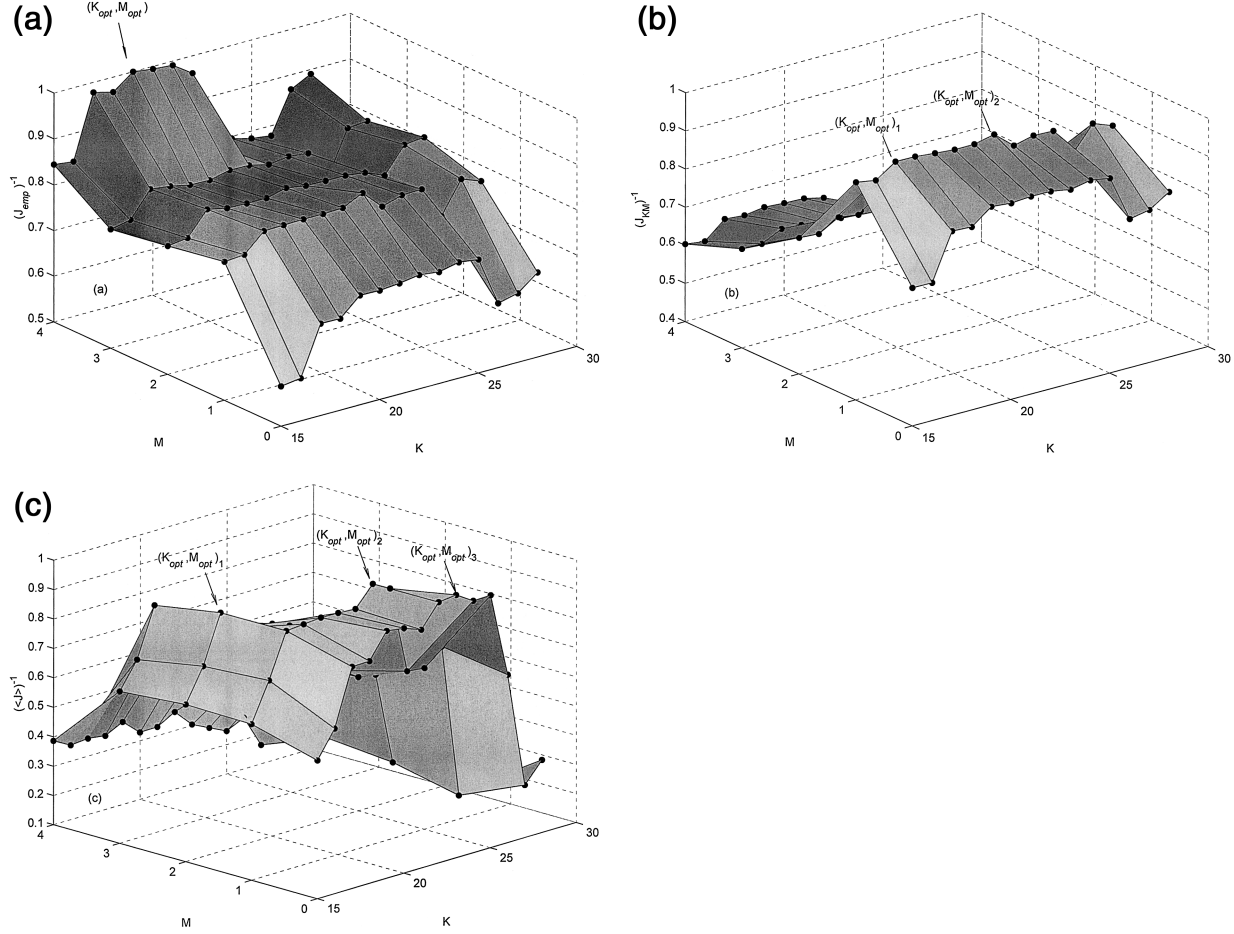
FIG. 4. Determination of optimal truncation of the toroidal and poloidal modes using (a) J_{emp}, (b) VCCF, and (c) $\langle J \rangle$.

2) VCCF

The dependence of VCCF on the mode truncation parameters K and M is demonstrated in Fig. 4b. Two maximum values of VCCF are found:

$$(K_{opt}, M_{opt})_1 = (17, 1), \quad \text{and} \quad (K_{opt}, M_{opt})_2 = (22, 1).$$

It is also noticed that the mode truncations varying from $(K_{opt}, M_{opt})_1$ to $(K_{opt}, M_{opt})_2$ does not change the VCCF noticeably.

3) STATISTICAL COST FUNCTION $\langle J \rangle$

Figure 4c shows the dependence of the statistical cost function $\langle J \rangle$ on the mode truncation parameters K and M. Three maximum values of $\langle J \rangle$ are found:

$$(K_{opt}, M_{opt})_1 = (17, 2), \quad (K_{opt}, M_{opt})_2 = (22,1), \quad \text{and}$$

$$(K_{opt}, M_{opt})_3 = (23, 0).$$

The statistical cost function $\langle J \rangle$ reaches the global extremum for $(K_{opt}, M_{opt})_3$. Notice that $\langle J \rangle$ is constructed from 1000 stochastic realizations of the distorted ob-

servations and, additionally, averaged over all grid points used for numerical circulation modeling.

f. Verification of reconstruction scheme for disposition A

We use the truncation

$$K_{opt} = 21, \quad M_{opt} = 1, \quad \text{and} \quad S_{opt} = 5 \quad (2.3)$$

to reconstruct the observational data with various stochastic noises.

1) WHITE NOISE ($\eta_1 = 0.3$)

A white noise with $\eta_1 = 0.3$ is added to the data collected from the disposition A (Fig. 2a). A typical distorted observational dataset for the reconstruction is illustrated in Fig. 3a.

(i) Zero-order approximation $\mathbf{h}^{[0]}$

Using the zero-order approximation $\mathbf{h}^{[0]}$, the reconstructed circulation (Fig. 5a) has a similar pattern as the
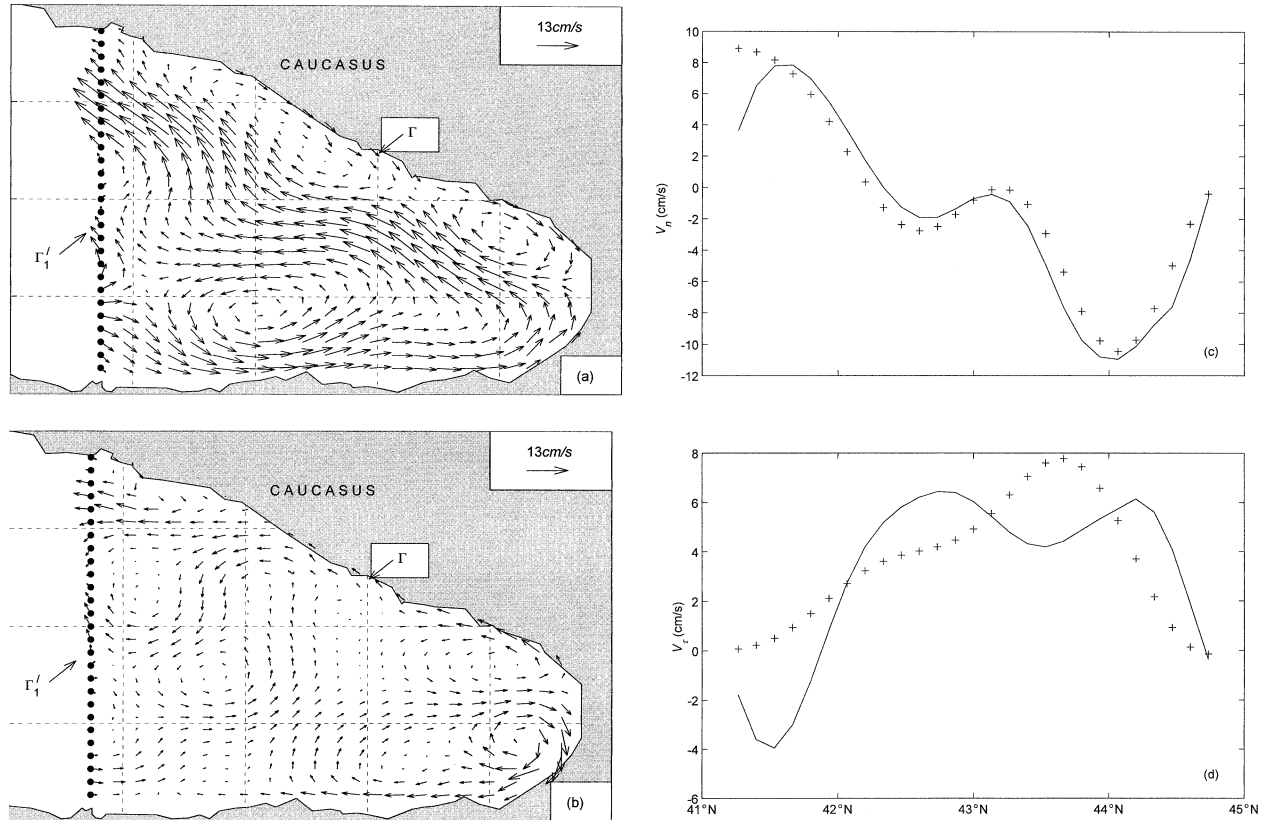
FIG. 5. Reconstructed velocity field from observed field indicated in Fig. 3a using the zero-order approximation: (a) velocity vectors, (b) residue velocity vectors, (c) normal velocity, and (d) tangential velocity at the open boundary. Here, the cross (+) and solid curve represent the reconstructed and exact velocity components at the open boundary.

exact field (Fig. 1): an areawide cyclonic gyre centered at 41.2°N, 39.15°E with a jet (rim current) around it. The center of the cyclone is shifted 20 km toward the southeast. The residue (exact minus estimated) velocity vector field (Fig. 5b) shows that the scheme overestimates the eastward jet near the southern boundary of the Black Sea and underestimates the westward jet in the northwestern part of the region. Besides, an anticyclone near the eastern corner of the Black Sea (centered at 42.7°N, 41.0°E, and 100 km in diameter) occurring in the reconstruct field (Fig. 5a) does not exist in the exact field (Fig. 1).

The capability of the reconstruction scheme is also recognized by the agreement between the reconstructed (cross) and exact (solid) normal (Fig. 5c) and tangential (Fig. 5d) velocity components at the open boundary. The reconstruction skill is represented by the root-mean-square (rms) errors [see (5.10) and (5.11) of Part I]:

$$\chi_u = 0.13, \qquad \chi_v = 0.18, \qquad \chi_n = 0.08, \quad \text{and}$$

$$\chi_\tau = 0.25. \tag{2.4}$$

Hence, the reconstruction quality is much higher for the normal component than for the tangential component.

(ii) Nonlinear regression

Using the nonlinear regression (see section 4b in Part I) with optimal truncations (2.3), the reconstructed circulation (Fig. 6a) has a pattern closer to the exact observational field (Fig. 1) than the zero-order approximation: a basinwide cyclonic eddy centered at 41°N, 39°E with a jet (rim current) and several small eddies around it. The center of this cyclone has the same location as the exact observational field. The residue velocity vector field (Fig. 6b) shows that the scheme underestimates the westward jet in the northwestern part of the region. Similar to the zero-order approximation, an anticyclone near the eastern corner of the Black Sea (centered at 42.5°N, 41.2°E, and 100 km in diameter) occurs in the reconstruct field (Fig. 6a), which does not exist in the exact observational field (Fig. 2). However, the strength of this false anticyclone is weaker than that in the zero-order approximation.

The reconstructed (cross) normal (Fig. 6c) and tangential (Fig. 6d) components agree well with the exact (solid) field at the open boundary. The reconstruction skill is represented by
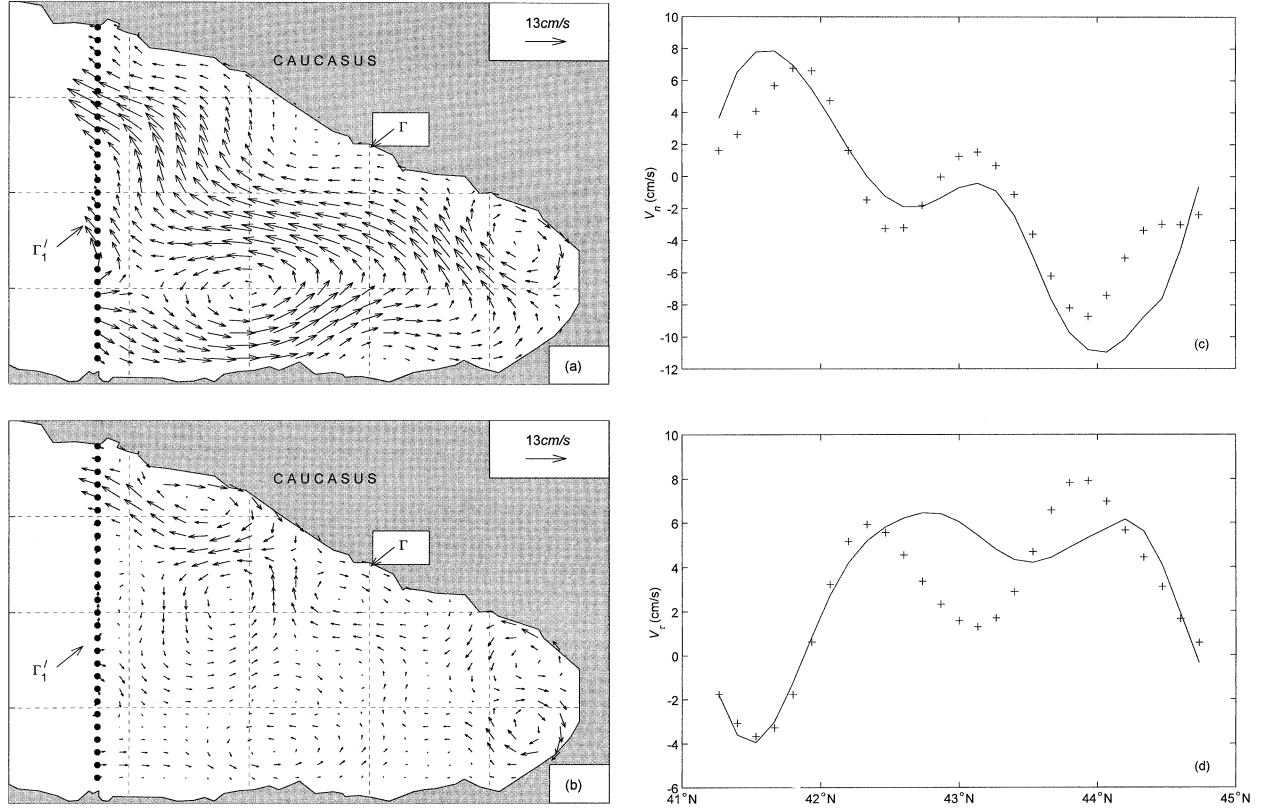
FIG. 6. Same as in Fig. 5, except using the nonlinear regression.

$$\chi_u = 0.10, \qquad \chi_v = 0.13, \qquad \chi_n = 0.11, \quad \text{and}$$

$$\chi_\tau = 0.14, \tag{2.5}$$

which indicates that the rms errors of all velocity components are less than or equal to 0.14. Comparison between (2.4) and (2.5) shows the improvement in reconstruction (except the normal velocity at the open boundary) using the nonlinear regression.

From theoretical point of view, the nonlinear regression should provide more accurate results than the zero-order approximation. However, comparison between (2.4) and (2.5) shows more accuracy on the normal velocity at the open boundary using the zero-order approximation rather than the nonlinear regression. This may be caused by strong dependence of the reconstruction accuracy on various factors such as station deposition pattern, level of noise in observations, and quality of numerical algorithm; further investigation is needed.

2) COLOR NOISE

A color noise with $\eta_1 = 0.3$, $L_{\text{cor}} = 45$ km is added to the data acquired from the disposition A (Fig. 2a). A typical distorted observational dataset for the reconstruction is illustrated in Fig. 3b.

(i) Zero-order approximation

Using the zero-order approximation $\mathbf{h}^{[0]}$, the reconstructed circulation (Fig. 7a) has a pattern similar to the exact circulation field (Fig. 1): a basinwide cyclonic gyre centered near 42°N, 39°E with a jet stream and several small eddies (about 100 km in diameter) around it, and this cyclone seems to be flattened. The east–west semiaxis is much longer than the north–south semiaxis (Fig. 7a). The residue velocity vector field (Fig. 7b) shows that the scheme overestimates both the eastward jet near the southern boundary and the westward jet in the northwestern part of the region. Note that an anticyclone near the upper half of the eastern corner of the Black Sea (centered at 42.75°N, 41°E, and 100 km in diameter) occurring in the reconstruct field (Fig. 7a) does not exist in the exact observational field (Fig. 1). At the open boundary, the reconstruction quality for the normal component (Fig. 7c) is much better than the tangential component (Fig. 7d). The reconstruction skill is represented by

$$\chi_u = 0.17, \qquad \chi_v = 0.16, \qquad \chi_n = 0.05, \quad \text{and}$$

$$\chi_\tau = 0.48, \tag{2.6}$$

which indicates that the reconstruction quality for the data with red noise is comparable to that for the data
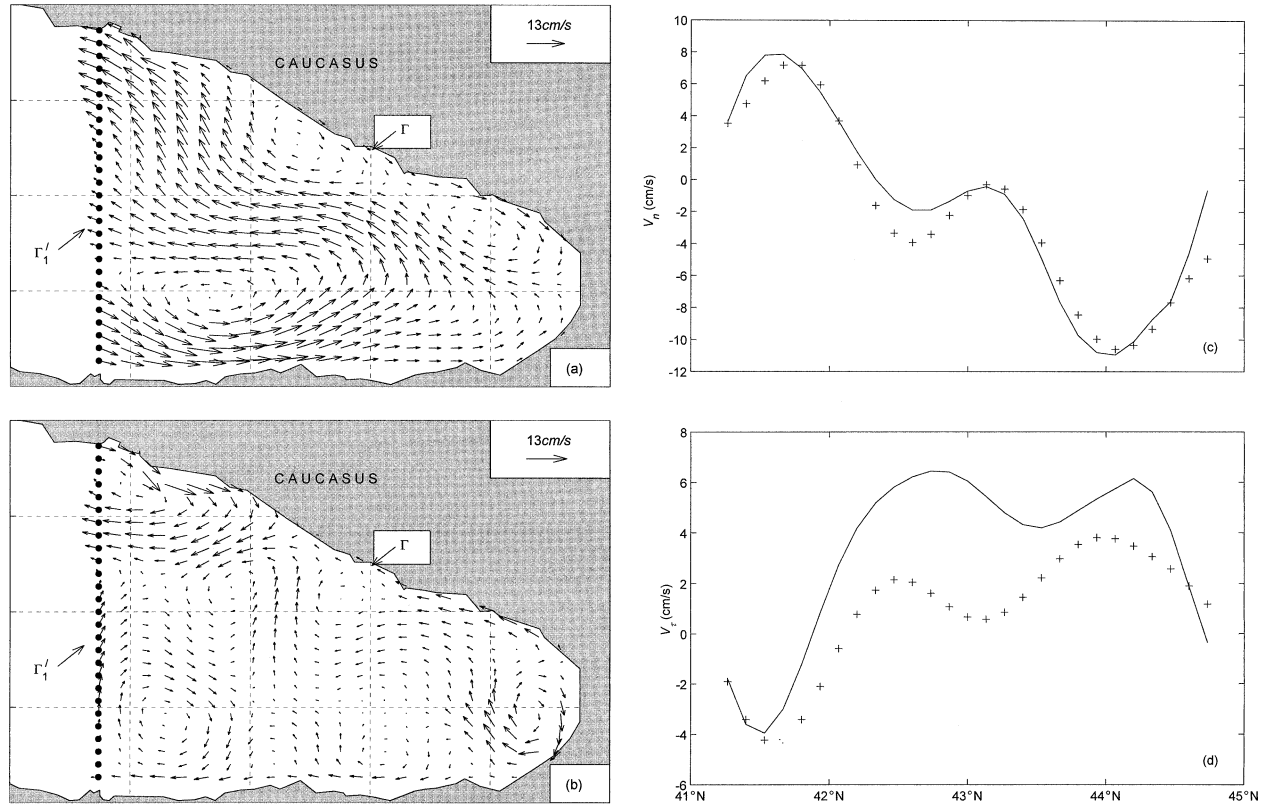
FIG. 7. Same as in Fig. 5, except from observed field indicated in Fig. 3b.

with white noise at the same noise-to-signal level (0.3) except for the tangential velocity at the open boundary. The tangential (normal) velocity error is much higher (lower) for the red noise than for the white noise.

(ii) Nonlinear regression

Using the nonlinear regression with optimal truncations (2.3), the reconstructed circulation (Fig. 8a) shows a similar pattern to the exact observational field (Fig. 1). The residue velocity vector field (Fig. 8b) shows that the scheme overestimates the eastward jet near the southern boundary and underestimates the westward jet in the northwestern part of the region. Similar to the zero-order approximation, an anticyclone near the eastern corner of the Black Sea (centered at 42.6°N, 41.2°E, and 100 km in diameter) occurs in the reconstruct field (Fig. 8a), which does not exist in the exact observational field (Fig. 1). At the open boundary, the reconstruction quality for the normal component (Fig. 8c) is comparable to that for the tangential component (Fig. 8d). The rms errors are computed by

$$\chi_u = 0.1, \qquad \chi_v = 0.29, \qquad \chi_n = 0.15, \quad \text{and}$$

$$\chi_\tau = 0.12. \tag{2.7}$$

As with the white noise, the nonlinear regression has less (more) accuracy than the zero-order approximation

in reconstructing the normal (tangential) component at the open boundary.

3) STRONG WHITE NOISE

To test the reconstruction capability in processing highly noisy data, a strong white noise ($\eta_1 = 1.2$) is added to the data acquired from the disposition A (Fig. 2a). The distorted observational dataset for the reconstruction is illustrated in Fig. 3c. The nonlinear regression with the optimal truncation (2.3) is used to reconstruct the velocity field.

The reconstructed circulation (Fig. 9a) still shows the existence of a basinwide cyclonic circulation. The major cyclone centered near 42°N, 39°E is also reconstructed. The residual velocity vector field (Fig. 9b) shows that the scheme overestimates the eastward jet near the southern boundary. The exact (solid) and reconstructed (cross) normal (Fig. 9c) and tangential (Fig. 9d) velocities at the open boundary show the capability of the scheme in reconstructing the velocity field at the open boundary. The reconstruction skill is represented by

$$\chi_u = 0.83, \qquad \chi_v = 1.22, \qquad \chi_n = 0.91, \quad \text{and}$$
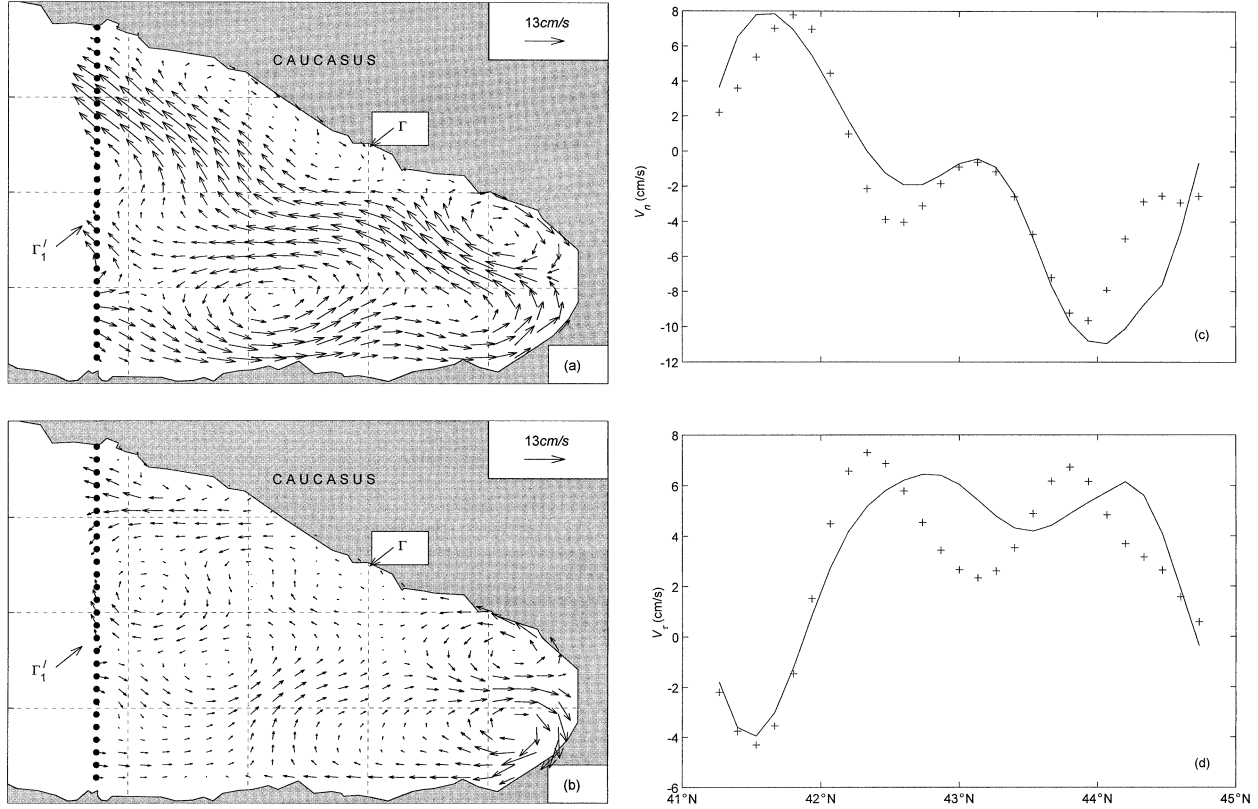
$$\chi_\tau = 0.83. \tag{2.8}$$

FIG. 8. Same as in Fig. 5, except from observed field indicated in Fig. 3b and using the nonlinear regression.

g. Sensitivity of reconstruction to sampling density

To test the dependence of the reconstruction capability on the sampling density, three observational datasets are acquired from dispositions B, C, and D, distorted by a white noise ($\eta_1 = 0.3$) (Fig. 3). The nonlinear regression is used to reconstruct the velocity fields. The rms errors (Table 1) show that the reconstruction quality deteriorates with the reduction of sampling density from high sampling density (disposition B, 19 observations) to medium sampling density (disposition C, 14 observations), and from medium sampling density to low sampling density (disposition D, 9 observations). For the low sampling density (disposition D), only six toroidal modes and one poloidal mode can be estimated. Any combination of these modes results in a poor reconstruction skill.

h. Reconstruction of vertical velocity

1) VERTICAL VELOCITY "DATA"

If surface gravity waves are ignored, the surface ($z = 0$) boundary condition may be written as

$$w = 0, \quad \text{at } z = 0. \tag{2.9}$$

The "exact" vertical velocity data (w_{ex}) at 20 m deep (Fig. 10a) are obtained for verification of the recon-struction scheme from vertical integration of the continuity equation of the Black Sea circulation model from the surface to 20 m.

2) TWO-STEP RECONSTRUCTION

The reconstruction process includes two steps. First, the observational horizontal velocity field is acquired from disposition A and distorted by a white noise ($\eta_1 = 0.3$) at $z = 0$, 5, 10, and 20 m, and the horizontal velocity is reconstructed at each level. Then the vertical gradient of poloidal potential (0 to 20 m) is determined. Notice that the vertical gradient of Φ does not equal zero at the ocean surface. Second, the integral equation (3.5) of Part I is solved to obtain the reconstructed vertical velocity w_{rec} at 20 m (Fig. 10b). Here, strong correlation is found between the vertical velocity (upwelling and downwelling zones; Fig. 10b) and the eddy structures in horizontal velocity vector field (Figs. 5a and 6a). The reconstruction error estimated by

$$w_{res} = w_{ex} - w_{rec}.$$

Large values of w_{res} (Fig. 10c) are found in some areas, because the reconstruction skill of the vertical velocity strongly depends on that of the horizontal velocity. At least five or six poloidal modes are needed to best resolve the horizontal velocity structure during recon-
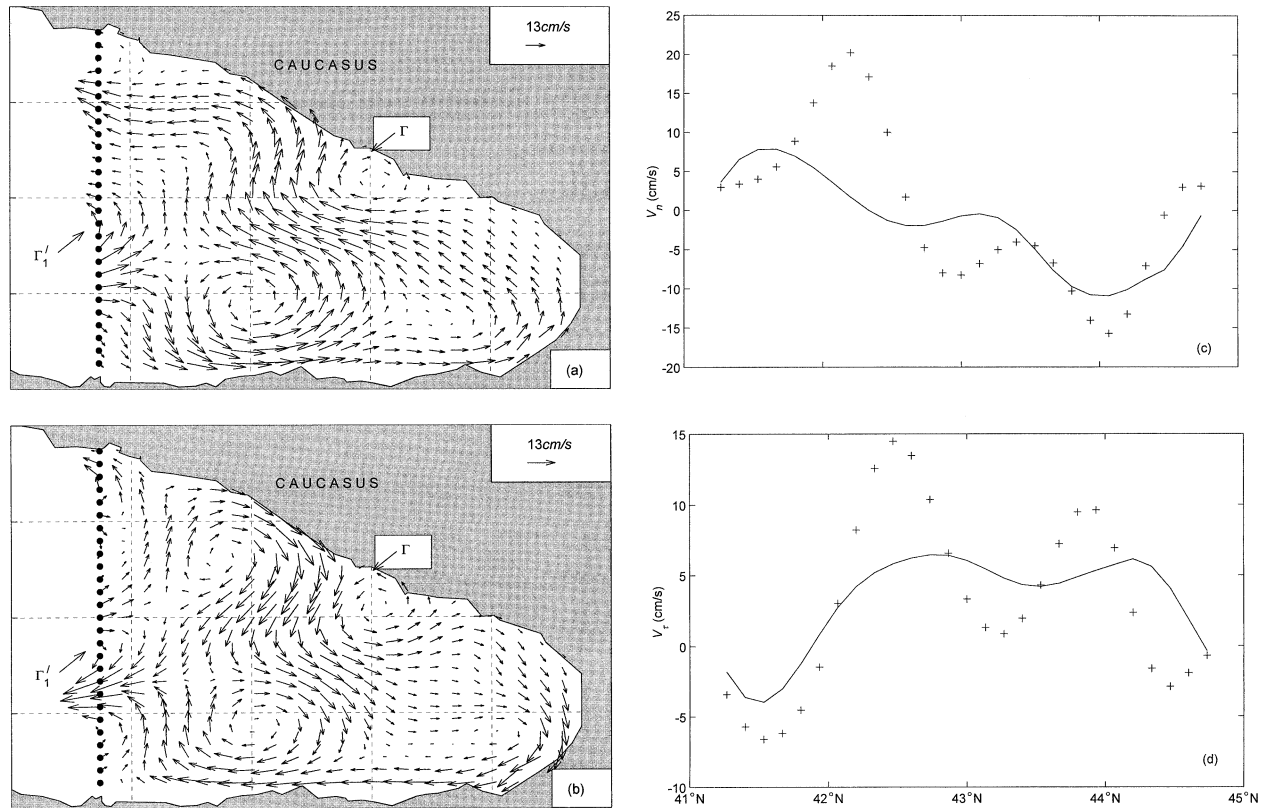
FIG. 9. Same as in Fig. 5, except from observed field indicated in Fig. 3c and using the nonlinear regression.

struction if the observational data contains no noise. For the selected station disposition and noise level, a maximum of two or three poloidal modes can be reconstructed. This limitation causes large difference between exact and reconstructed vertical velocity in some areas.

i. Comparison with the smoothed splines

The reconstructed field is compared to that reproduced using the smoothed splines whose smoothing parameter is obtained using Tikhonov's regularization procedure. Both techniques weakly depend on the noise level in observations if $\eta_1 \leq 0.3$. For example, the reconstruction skill of the smoothed splines over the observational data acquired from disposition A and distorted by a white noise ($\eta_1 = 0.3$), is given by

$$\chi_u = 0.15, \qquad \chi_v = 0.23, \qquad \chi_n = 0.18, \quad \text{and}$$

$$\chi_\tau = 0.43, \tag{2.10}$$

TABLE 1. Reconstruction skill for various disposition patterns.

Sampling pattern	X_u	X_v	X_n	X_τ
Disposition B	0.30	0.23	0.18	0.43
Disposition C	0.28	0.36	0.49	0.48
Disposition D	0.32	0.40	0.55	0.62

which is comparable to (2.4) and (2.5). For the observational data collected from disposition A and distorted by a strong white noise ($\eta_1 = 1.2$), the rms errors are extremely high using the smooth splines:

$$\chi_u = 1.18, \qquad \chi_v = 1.73, \qquad \chi_n = 0.65, \quad \text{and}$$

$$\chi_\tau = 1.1. \tag{2.11}$$

Comparison between (2.11) and (2.8) confirms the capability of the reconstruction scheme using the two potential scalar representation.

3. Case 2: Shallow water flow in a semienclosed basin

a. Geometry of the basin

A shallow water basin with a flat bottom is centered at 35°N and bounded by three rigid boundaries. This basin extends 1000 km in the north–south and 1050 km in the east–west directions. The northern, southern, and western boundaries are rigid, and the eastern boundary is open. The Cartesian coordinate system is chosen with the origin at the southwest corner. The x axis points toward the east, and the y axis toward the north (Fig. 11a).
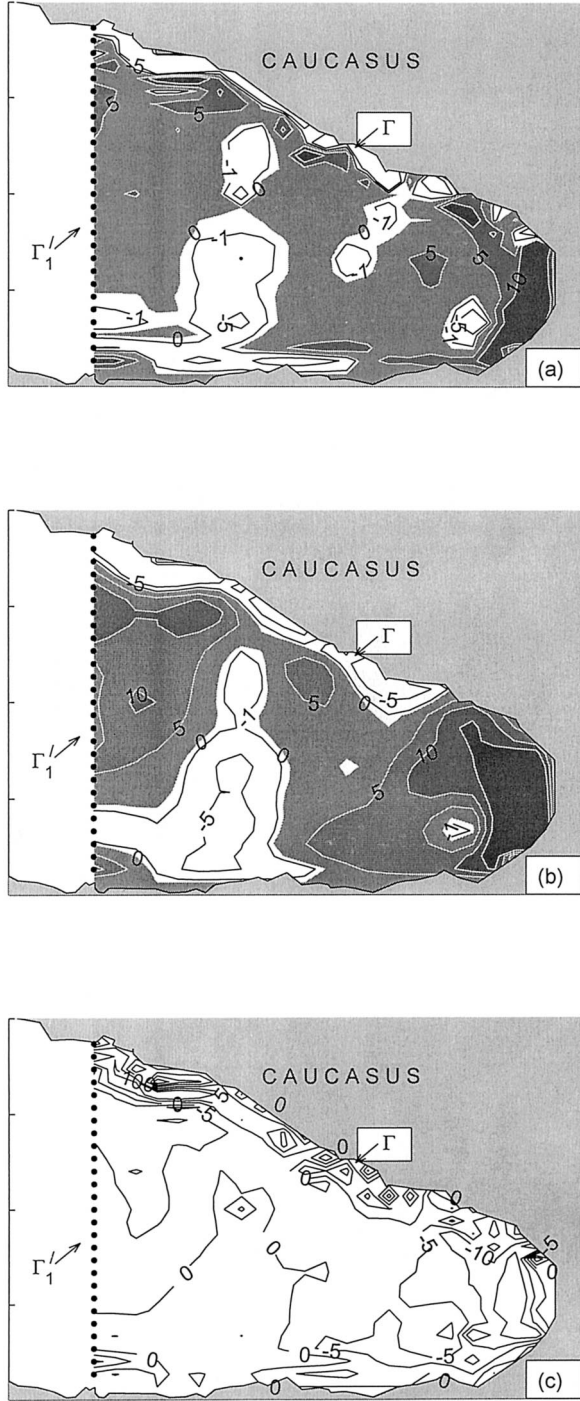
FIG. 10. Vertical velocity (in 10^{-4} cm s^{-1}) at 20 m deep: (a) exact data, (b) reconstructed field from depostion A sampling with a white noise ($\eta_1 = 0.3$), and (c) error (w_{res}). The positive values indicate upwelling.

b. Exact velocity data

The exact velocity data for the semienclosed basin is model generated with the Princeton Ocean Model (POM) developed by Blumberg and Mellor (1987).
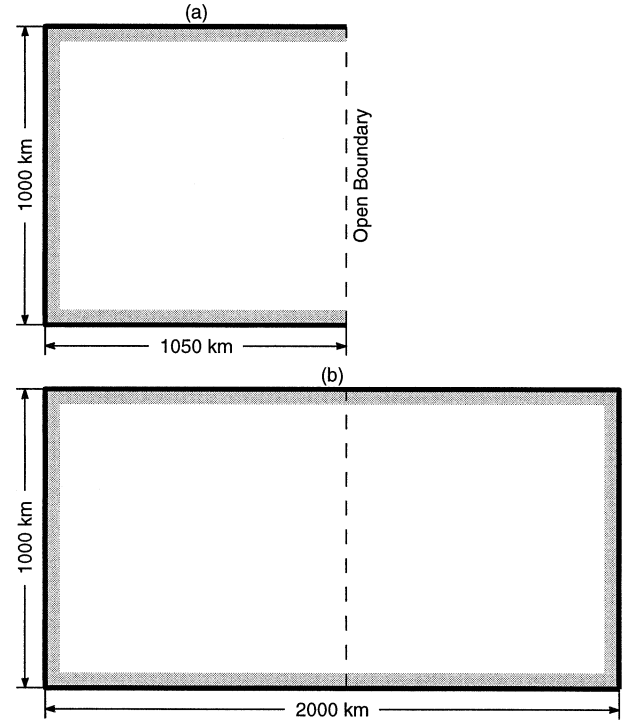

FIG. 11. Two areas for the POM integration: (a) domain with three rigid and one open boundaries, and (b) a domain with four rigid boundaries.

POM is a primitive equation model with a free surface and a level 2 1/2 turbulence closure scheme. We use the 2D version of POM to get the exact data (Chu et al. 1997) for verifying the reconstruction scheme. The eastern boundary is connected to a mirror image of domain A (about $x = 1050$ km), forming a closed rectangular domain (Fig. 11b). The POM model was integrated for the large domain with four rigid boundaries from the following initial conditions:

$$(u, \, v, \, w) = 0 \text{ (in m s}^{-1}\text{)},$$

$$T = 283\left[1 + \exp\left(\frac{z}{1000}\right)\right] \text{ (in K)}, \quad \text{and}$$

$$S = 35 \text{ (in ppt)},$$

and with no surface heat or salinity fluxes but zonal surface pseudo–wind stress varying with latitude:

$$\frac{\tau_x}{\rho_0} = -10^{-3} \cos\frac{\pi y}{L_y}, \qquad L_y = 1000 \text{ km.} \quad (3.1)$$

The model time step was 2.5 min. The horizontal resolution is 50 km. Bottom stress is parameterized by the quadratic drag relation and the horizontal kinematic viscosity is set to 0. When POM is integrated from the rest ($u = v = 0$) and the flat sea surface ($\xi = 0$), two circulation regimes are identified in the basin: (i) a linear regime (up to day 20) with velocities under 10 to 15 cm s^{-1} and surface sea elevation reaching 5–10 cm; (ii)

(a)

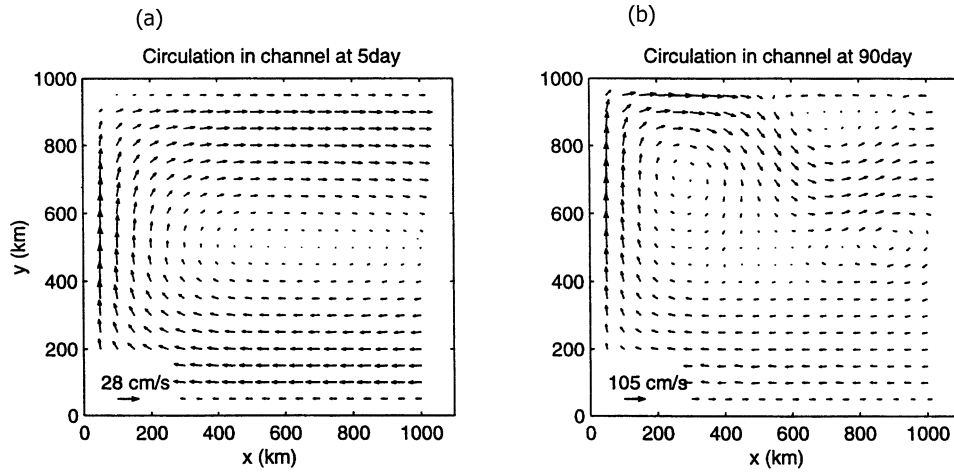Circulation in channel at 5day

(b)

Circulation in channel at 90day

FIG. 12. Exact velocity dataset for the semienclosed basin: (a) day 5 and (b) day 90.

a nonlinear regime (from day 20 to day 50) with occurrence of large velocities (100 to 120 cm s^{-1}) and surface elevation (100 cm). For the nonlinear regime, we have

$$\sigma_1 = \max\left(\frac{U}{fL_\perp}\right) \sim 0.25, \qquad \sigma_2 = \max\left(\frac{1}{fT_\perp}\right) \sim 0.2,$$

where $U_\perp$ is the characteristic velocity in the basin; $L_\perp$ and $T_\perp$ are the spatial and temporal scales of the flow, respectively; f is the Coriolis parameter. The POM-generated 2D velocity fields on day 5 (in the linear regime; Fig. 12a) and day 90 (in the nonlinear regime; Fig. 12b) are treated as the exact velocity data.

c. Sampling strategies

The following six sampling strategies (Fig. 13) are used to investigate the effect of data distribution on the quality of reconstruction: (a) irregular basinwide disposition with high density (86 points), (b) irregular basinwide disposition with lower density (42 points), (c) regular basinwide disposition (192 points), (d) irregular partial domain (noncoastal) disposition (69 points), (e) regular coastal disposition (114 points), and (f) irregular coastal disposition (69 points). Notice that there are no observations at the open boundary. Here, we are particularly interested in the reconstruction quality for partial sampling cases (i.e., dispositions D, E, and F), where the observational domain (S) is smaller than the reconstruction domain (S'), that is $S \subset S'$.

d. Simplification of two-scalar representations

For a 2D shallow water flow, the toroidal and poloidal potentials coincide with the streamfunction and velocity potential (Chu 1999; Chu et al. 2003). The velocity (u, v) and surface elevation (ξ) are represented by

$$u = \frac{1}{H + \xi}\left(\frac{\partial\Psi}{\partial y} + \frac{\partial\Phi}{\partial x}\right),$$

$$v = \frac{1}{H + \xi}\left(-\frac{\partial\Psi}{\partial x} + \frac{\partial\Phi}{\partial y}\right), \quad \text{and}$$

$$\frac{\partial\xi}{\partial t} = -\Delta\Phi, \qquad (3.2)$$

and the streamfunction Ψ satisfies the Poisson equation

$$\Delta\Psi = \frac{\partial\xi}{\partial y}u - \frac{\partial\xi}{\partial x}v - (H + \xi)\zeta, \qquad (3.3)$$

where ζ is vertical vorticity.

Scale analysis leads to

$$\frac{|\xi|}{H} \simeq O(10^{-4}), \qquad \frac{|u\partial\xi/\partial y - v\partial\xi/\partial x|}{|(H + \xi)\zeta|} \simeq O(10^{-4}),$$

$$\frac{|\partial\Phi/\partial y|}{\partial\Psi/\partial x|} \simeq O(10^{-6}), \qquad \frac{|\partial\Phi/\partial x|}{|\partial\Psi/\partial y|} \simeq O(10^{-6}).$$

Thus, Eq. (3.2) is simplified into

$$u \simeq \frac{1}{H}\frac{\partial\Psi}{\partial y}, \qquad v \simeq -\frac{1}{H}\frac{\partial\Psi}{\partial x}. \qquad (3.4)$$

Thus, the reconstruction is reduced to the analysis of one scalar potential (i.e., the streamfunction Ψ). Notice that the neglect of the second scalar potential (Φ) is accurate only for the selected rectangular basin characteristics. The exact data illustrated in Figs. 12a and 12b show that at the open boundary the normal component is much larger than the tangential component $\oint_L |V_n|\ d\tau \gg \oint_L |V_\tau|\ d\tau$ at open boundary. We may neglect V_τ and use $\kappa = 0$ at the open boundary [see (2.10) of Part I] for the reconstruction.
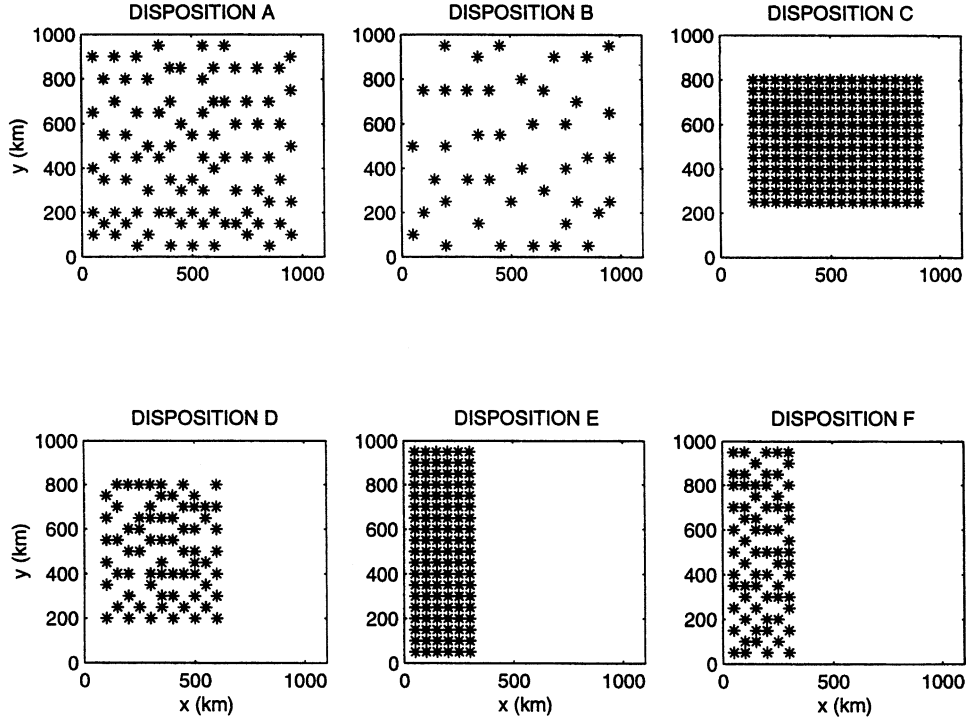
FIG. 13. Several sampling strategies: (a) irregular domainwide disposition with high density (86 points), (b) irregular domainwide disposition with lower density (42 points), (c) regular partial domainwide disposition (192 points), (d) irregular partial domain (noncoastal) disposition (69 points), (e) regular coastal disposition (114 points), and (f) irregular coastal disposition (69 points).

e. Noise reduction using the Gram–Schmidt process

The basis functions $\{\Psi_k \mid k = 0, 1, \ldots, K\}$ are calculated using MATLAB (1997), and estimate the spectral coefficients a_k using the linear regression model. The optimal mode truncation is obtained using VCCF [see (3.3) in Part I]. The truncation error

$$\Psi_{\text{pertub}} = \sum_{K}^{\infty} a_k \Psi_k$$

is defined as the deterministic perturbation. The noise-to-signal ratio associated with the deterministic perturbation can reach 0.1. Such a level of noise may considerably deteriorate the reconstruction skill, especially for a partially sampled region such as dispositions C, D, and E (Fig. 13).

The regularization procedure [see (4.6)–(4.9) of Part I] effectively filters noise if vectors $\mathbf{Y}_{\text{noise}}$ and $\mathbf{Y}_{\text{exact}}$ are perpendicular:

$$\mathbf{Y}_{\text{noise}} \cdot \mathbf{Y}_{\text{exact}} = 0. \tag{3.5}$$

Therefore, the reconstruction procedure is transformed to satisfy (3.5), which can be done through selecting appropriate basis functions $\{\Psi_k\}$.

Since the number of observations (Fig. 13) is sufficiently large, two sets of basis functions can be constructed from the gradients of Ψ_k:

$$\Theta_k = \frac{\partial \Psi_k}{\partial y} \quad \text{and} \quad \Xi_k = \frac{\partial \Psi_k}{\partial x},$$

which can be orthonormalized using the Gram–Schmidt process within S. Thus, we represent each of the velocity components as

$$u_K = \sum_{k=0}^{K} a_k \Theta_k \quad \text{and} \quad v_K = -\sum_{k=0}^{K} b_k \Xi_k. \tag{3.6}$$

Next, the two sets $\{\Theta_k\}$ and $\{\Xi_k\}$ are reorthonormalized within the domain S' into

$$(H \Theta_k' \, \Theta_m') = \delta_{km}, \qquad (H \Xi_k' \, \Xi_m') = \delta_{km},$$

where Θ_m' and Ξ_m' are renormalized basis functions,

$$H = \begin{cases} 1, & x \in S \\ 0, & x \notin S \end{cases}$$

is the weight, and δ_{nm} is the Kroneker delta. The velocity components can be expanded into

$$u_N = \sum_{m=0}^{K} A_m \Theta_m' \quad \text{and} \quad v_N = -\sum_{m=0}^{K} B_m \Xi_m', \tag{3.7}$$

and new spectral coefficients A_m and B_m are computed by

$$A_m = \sum_{k=0}^{K} \mu_{mk}^{(1)} a_k \quad \text{and} \quad B_m = \sum_{k=0}^{K} \mu_{mk}^{(2)} b_k, \tag{3.8}$$

where $\mu_{mk}^{(1)}$ and $\mu_{mk}^{(2)}$ are the coefficients generated using the Gram–Schmidt process. Using new basis functions $\{\Theta_m'\}$, $\{\Xi_m'\}$, the coefficients A_m and B_m are estimated from the observations inside S by the simplest regression model. The accuracy of the reconstruction should be improved since the influence of deterministic perturbation on the coefficients has been reduced. Then, the spectral coefficients a_k and b_k are calculated from (3.8), which gives an approximate solution for the domain S'.

f. Reconstructed velocity field

Since the semienclosed basin (Fig. 11) is rectangular and $\kappa = 0$ is used for the open boundary condition [see (2.10) of Part I], the basis functions Ψ_k may be replaced by the trigonometric polynomials. Such a replacement may simplify the theoretical analysis; however, the convergence of the spectral expansion of the streamfunction is much slower using the trigonometric polynomials than using the basis functions. Thus, even for a basin with simplified geometry (such as a rectangle) basis functions are still preferred.

1) DETERMINATION OF THE OPTIMAL MODE NUMBER K_{opt}

The optimal mode number K_{opt} strongly depends on the sampling strategy and the level of noise in the observation. To illustrate this, the day 5 observational data are acquired from disposition B (Fig. 13b) with a red noise ($\eta_1 = 0.5$, $R_{cor} = 70$ km) and the day 90 observational data are acquired from disposition A (Fig. 13a) with a white noise ($\eta_1 = 0.5$). Using VCCF, the optimal truncation K_{opt} is determined as 25–30 (50 to 75) for day 5 (day 90) circulation.

Taking the regularization parameter δ [see (4.15) in Part I] as 0.1, the noisy data are reconstructed on day 5 (Fig. 14a) and day 90 (Fig. 14c). Comparison between Figs. 14a and 12a (Figs. 14c and 12b) shows that the reconstruction scheme reproduces almost all the details of circulation pattern. The residual velocity vector fields for day 5 and day 90 are represented in Figs. 14b and 14d, respectively. The reconstruction scheme underpredicts the basinwide circulation for day 5.

2) EFFECT OF NOISE ON THE RECONSTRUCTION SKILL

To test the sensitivity of reconstruction skill on the noise level and characteristics, six different types of noise are added to the exact observational data with various noise levels ($0 \leq \eta_1 \leq 1.0$) on day 90: (a) white noise, (b) color noise with $R_{cor} = 7.5$ km, (c) color noise with $R_{cor} = 70$ km, (d) color noise with $R_{cor} = 100$ km, (e) color noise with $R_{cor} = 200$ km, and (f) color noise with $R_{cor} = 500$ km. The velocity field is reconstructed using the schemes presented in Part I, and the rms errors of u component (Fig. 15) and v component (Fig. 16)
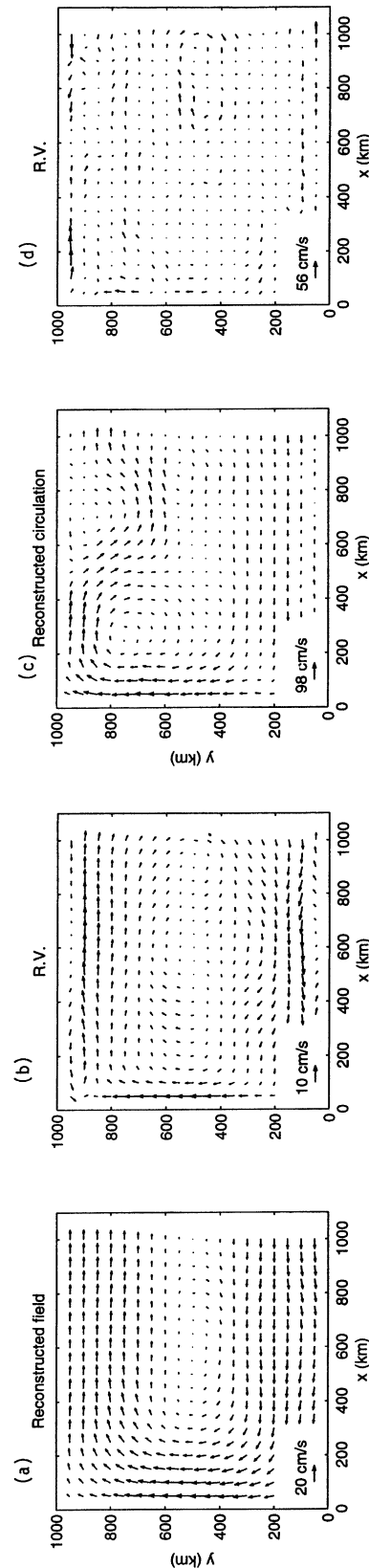

FIG. 14. (a) Velocity field and (b) residue velocity on day 5 reconstructed ($\delta = 0.1$ and $K_{opt} = 30$) from observed data collected from disposition B with a color noise $\eta_1 = 0.5$, $R_{cor} = 70$ km; (c) velocity field and (c) residue velocity on day 90 reconstructed ($\delta = 0.1$ and $K_{opt} = 50$) from observed data collected from disposition A with a white noise with $\eta_1 = 0.5$.

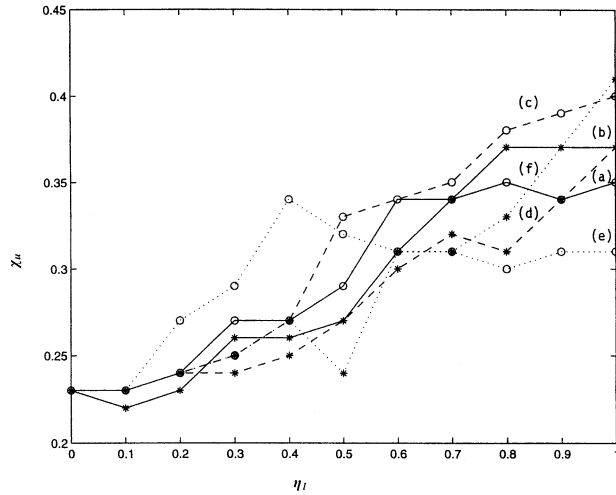FIG. 15. Dependence of rms errors for u component (on day 90) on η_1 for different noise characteristics: (a) white noise, (b) color noise with $R_{cor} = 7.5$ km, (c) color noise with $R_{cor} = 70$ km, (d) color noise with $R_{cor} = 100$ km, (e) color noise with $R_{cor} = 200$ km, and (f) color noise with $R_{cor} = 500$ km.

FIG. 16. Same as in Fig. 15, except for v component.

are computed for different noise characteristics. The results show that the reconstruction skill is quite high ($\chi_u < 0.3$ and $\chi_v < 0.2$) and weakly sensitive to η_1 and to correlation radius of color noise if $\eta_1 \leq 0.5$ and $R_{cor} \leq 500$ km (Figs. 15 and 16).

3) NORMAL VELOCITY AT THE OPEN BOUNDARY

To verify the reconstruction quality at the open boundary, the reconstructed (denoted by +) and exact normal (solid curve) velocities at the open boundary are presented for various noise characteristics (noise level and correlation scale for color noise) and sampling strategies: disposition A (Fig. 17), disposition B (Fig. 18), and disposition D (Fig. 19). Among the six patterns (Fig. 13), dispositions A, B, and C are domainwide sampling (some observations close to the open boundary); disposition D is a partial domain sampling; and dispositions E and F are coastal sampling (all observations far away from the open boundary).

(i) Effect of sampling

For domainwide sampling (Figs. 17 and 18), the normal velocity is reproduced quite well for evident noises ($\eta_1 = 0.5$). However, for partial domain sampling (Fig. 19), the normal velocity is reproduced only for the disposition D with weak ($\eta_1 = 0.1$) white noise, and not for the coastal sampling (dispositions E and F) where the observations are far from the open boundary. Such a failure is due to the fact that the developed noise reduction procedure (3.6)–(3.8) cannot totally eliminate the noise caused by the truncation error.

Generally, the reconstruction skill does not directly depend on the number of observations, but does depend on the conditional number of the coefficient matrix **A**
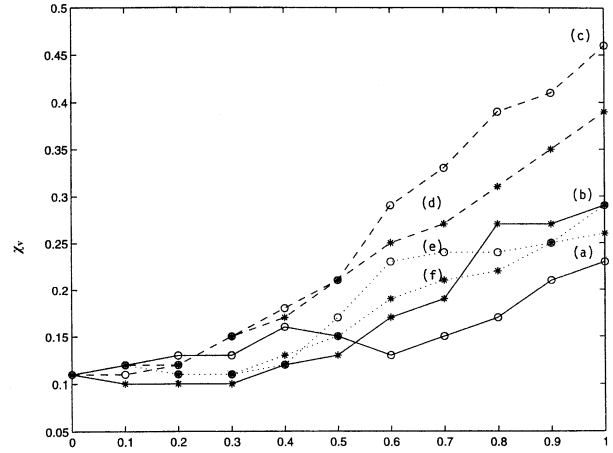
of the linear algebraic equation [(4.5) in Part I]. This number is determined by the station disposition pattern. The farther the observations are to the open boundary, the larger the conditional number will be. It reaches 10^8 for dispositions E and F. For a high conditional number, a weak noise in the observation will destroy the reconstruction. To improve this, the approach should be combined with a hydrodynamic model, which is the ongoing project.

(ii) Effect of deterministic disturbance for domainwide sampling

Without stochastic noise (Figs. 17 a,b; Fig. 18a), the difference between reconstructed and exact normal velocities shows the effect of a deterministic disturbance (i.e., the effect of the mode truncation) is caused by the truncation error. For observations sampled on day 90 with disposition A (Fig. 17), the reconstruction skill is almost the same between a mode truncation of 50 (Fig. 17a) and 70 (Fig. 17b), which indicates that the reconstruction skill is not sensitive to K_{opt} within the range (50–75 for day 90) determined using VCCF.

With a color noise ($\eta_1 = 0.5$, $R_{cor} = 70$ km) in observations sampled on day 5 with disposition B (Figs. 18b,d) the reconstruction skill without regularization ($\delta = 0$) is lower with $K_{opt} = 21$ (Fig. 18d) than with $K_{opt} = 30$, which indicates that the reconstruction skill is sensitive to K_{opt} when it varies out of the range (25–30 for day 5) determined using VCCF.

(iii) Effect of stochastic noise characteristics for domainwide sampling

The reconstruction skill is not sensitive to noise characteristics if the mode truncation K_{opt} varies within the range (25–30 for day 5, and 50–75 for day 90), determined by VCCF (all the panels in Figs. 17 and 18 except Fig. 18d). The small difference between Figs. 17c,e,g, and Figs. 17d,f,h shows that the regularization ($\delta \neq 0$) does not significantly improve the reconstruction skill
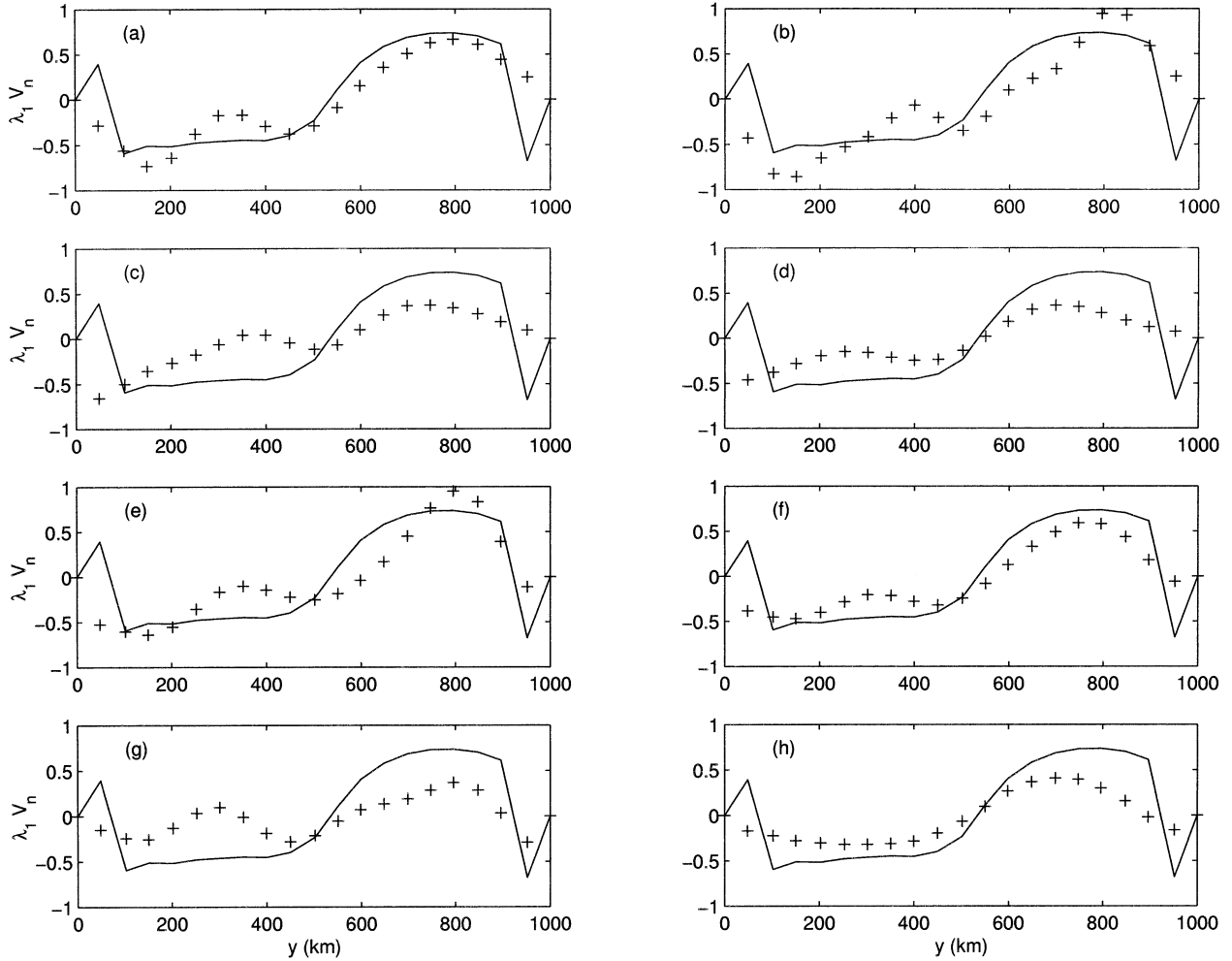
FIG. 17. Reconstructed (from disposition A, denoted by "+") and exact (solid curve) normal velocities at the open boundary on day 90 (a) without stochastic noise $K_{opt} = 50$; (b) without stochastic noise $K_{opt} = 70$; (c) with a white noise $\eta_1 = 0.5$, and $K_{opt} = 56$, $\delta = 0$; (d) with a white noise $\eta_1 = 0.5$, and $K_{opt} = 56$, $\delta = 0.1$; (e) with a color noise $\eta_1 = 0.5$, $R_{cor} = 100$ km, and $K_{opt} = 50$, $\delta = 0$; (f) with a color noise $\eta_1 = 0.5$, $R_{cor} = 100$ km, and $K_{opt} = 50$, $\delta = 0.1$; (g) with a color noise $\eta_1 = 0.5$, $R_{cor} = 200$ km, and $K_{opt} = 50$, $\delta = 0$; (h) with a color noise, $\eta_1 = 0.5$, $R_{cor} = 200$ km, and $K_{opt} = 50$, $\delta = 0.1$.

without regularization ($\delta = 0$) for domainwide sampling with high density. Notice that the mode truncation K_{opt} in all cases illustrated in Fig. 17 is within the range (50–75 for day 90) determined using VCCF.

4. Case 3: Reconstruction from Lagrangian drifter data in the Southern Ocean

The large-scale circulation is reconstructed from the observations of more than 200 surface drifting buoys deployed during FGGE between 1 December 1978 and 1 December 1979 to illustrate the capability of the scheme on processing the Lagrangian-type data.

a. FGGE data

The original buoy data consist of nonuniform time series of position, sea level pressure, and sea surface

temperature. Garrett (1980) discussed the accuracy of the measurements and buoy data acquisition system. The long-term trajectories of these buoys in the Southern Ocean and surrounding waters (south of 25°S) are examined (Fig. 20).

The surface velocity from a drifter buoy $\mathbf{V}_{drift}$ can be represented by

$$\mathbf{V}_{drift} = \mathbf{V}_{geostr} + \mathbf{V}_{Ekm} + \mathbf{V}_{Stok} + \mathbf{V}_{drag}, \qquad (4.1)$$

where $\mathbf{V}_{geostr}$, $\mathbf{V}_{Ekm}$, $\mathbf{V}_{Stok}$, and $\mathbf{V}_{drag}$ are the geostrophic, Ekman, Stokes, and drag velocities, respectively. Daniault and Menard (1985) estimated that the mean surface velocity $\langle \mathbf{V}_{drift} \rangle$ over the Southern Ocean is about 24 cm s^{-1}; and the summation of the Ekman, Stokes, and drag velocities (i.e., the ageostrophic velocity) is around 7 cm s^{-1}. Therefore, the ageostrophic component is significantly large (around 30%) in the drifter buoy data.
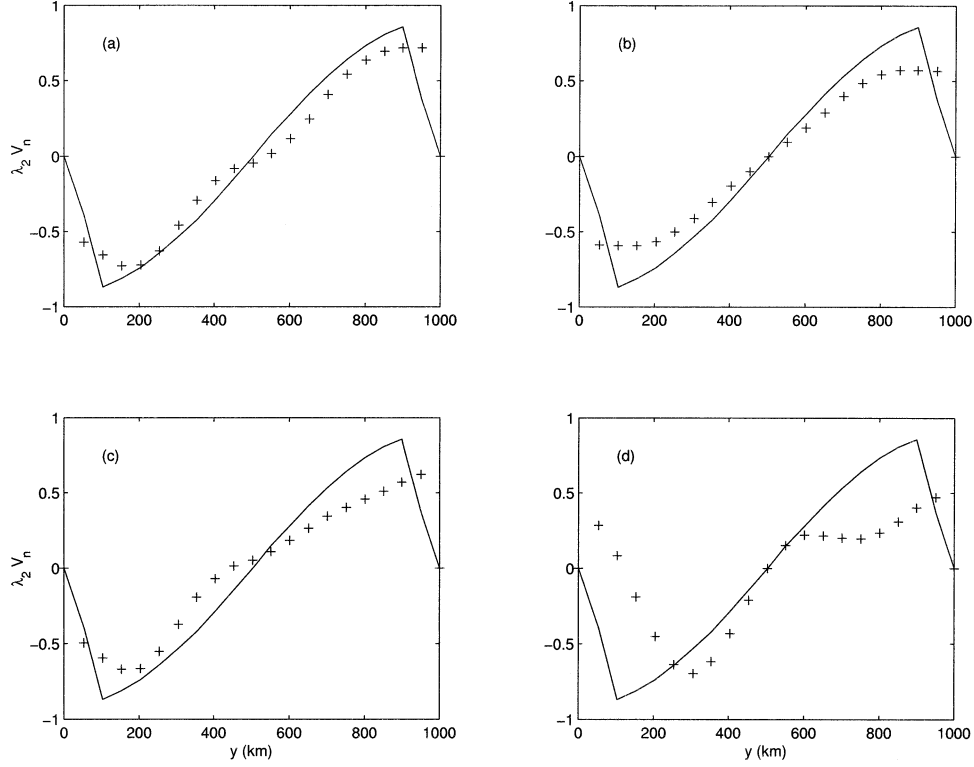
FIG. 18. Reconstructed (from disposition B, denoted by "+") and exact (solid curve) normal velocities at the open boundary on day 5 (a) without stochastic noise $K_{opt} = 30$; (b) with a color noise $\eta_1 = 0.5$, $R_{cor} = 70$ km, and $K_{opt} = 30$, $\delta = 0$; (c) with a color noise $\eta_1 = 0.5$, $R_{cor} = 70$ km, and $K_{opt} = 30$, $\delta = 0.1$; (d) with a color noise $\eta_1 = 0.5$, $R_{cor} = 70$ km, and $K_{opt} = 21$, $\delta = 0$.

Thus, both scalar potentials Ψ and Φ are used to reconstruct the data.

b. Data preparation

The trajectories of the drifter buoys were smoothed using 2D splines to exclude any spatial peculiarities along the trajectories, which cannot be resolved on a $2° \times 2°$ grid. Figure 21 shows the Lagrangian velocities along the smoothed trajectories for May 1979. After the smoothing, the ratio between the filtered (E_{filtr}) and total (E_{tot}) kinetic energies is given by

$$\frac{E_{filtr}}{E_{tot}} \simeq 0.2 - 0.25.$$

c. Domains A and B

We reconstruct the circulation separately in two areas: domains A and B. The domain A (Fig. 22a) is enclosed by three rigid boundaries: Γ (70°S); Γ_2 (25°S and connected coastlines of South America, South Africa, Madcascar, and Australia south of 25°S); and Γ_1 (representing New Zealand). Notice that the boundary conditions for domain A are often used in large-scale ocean circulation modeling for the Southern Ocean.

The domain B (Fig. 22b) is enclosed by two rigid boundaries: Γ and Γ_3 (coastline of South America south of 48°S) and an open boundary Γ' (48°S latitude over water). Comparison between the two reconstructed velocity fields leads to the verification of the open boundary condition [(2.10) of Part I] and the boundary extension method (section 4e of Part I).

d. Spectral decomposition

Both domains are multiply connected. Here, domain A has two islands (enclosed by Γ and Γ_1), and domain B has one island (enclosed by Γ), respectively. Following section 2d(3) of Part I, the toroidal and poloidal spectral expansions for domains A and B are represented by

$$\Psi^A = C_1 \Psi_{harm}^{(1)} + C_2 \Psi_{harm}^{(2)} + \sum_{k=0}^{K} a_k \Psi_k^A,$$

$$\Psi^B = C_3 \Psi_{harm}^{(3)} + \sum_{k=0}^{K} a_k \Psi_k^B, \qquad (4.2)$$

$$\frac{\partial \Phi^A}{\partial z} = \sum_{m=0}^{M} b_m \Phi_m^A, \qquad \frac{\partial \Phi^B}{\partial z} = \sum_{m=0}^{M} b_m \Phi_m^B, \quad (4.3)$$

where (Ψ_k^A, Ψ_k^B) and (Φ_m^A, Φ_m^B) are the eigenfunctions of the Laplacian operator calculated for the two domains
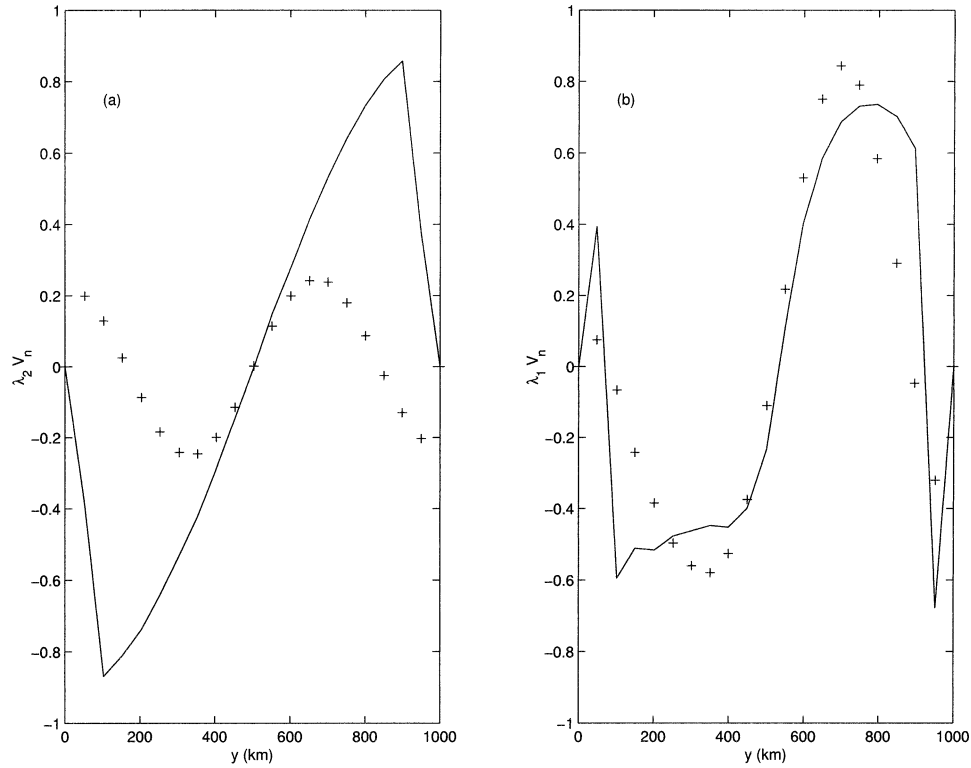
FIG. 19. Reconstructed (from disposition D, denoted by "+") and exact (solid curve) normal velocities at the open boundary white a noise ($\eta_1 = 0.1$) on (a) day 5 with $K_{opt} = 28$, and (b) day 90 with $K_{opt} = 75$.

with homogeneous boundary conditions [(2.24), (2.27), and (2.28) in Part I]. The additional three harmonic functions $\Psi_{harm}^{(1)}$, $\Psi_{harm}^{(2)}$, and $\Psi_{harm}^{(3)}$ satisfy the Laplacian equation with the boundary conditions written by

$$\Psi_{harm}^{(1)}|_\Gamma = 0, \qquad \Psi_{harm}^{(1)}|_{\Gamma_1} = 1, \qquad \Psi_{harm}^{(1)}|_{\Gamma_2} = 0,$$

$$\Psi_{harm}^{(2)}|_\Gamma = 0, \qquad \Psi_{harm}^{(2)}|_{\Gamma_2} = 1, \qquad \Psi_{harm}^{(1)}|_{\Gamma_1} = 0,$$

$$\Psi_{harm}^{(3)}|_\Gamma = 0, \qquad \Psi_{harm}^{(3)}|_{\Gamma_3 \cup \Gamma'} = 1. \qquad (4.4)$$

FIG. 20. Original FGGE buoy trajectories in the Southern Ocean.

e. Determination of coefficients

The spectral coefficients C_1, C_2, and C_3 can be determined from data or using special integral relationships (Kamenkovick 1961; McWilliams 1977; Flierl 1977; Platzman 1979) around any island:

$$\oint_\Lambda \frac{\partial \Psi}{\partial n} d\tau' = F_\tau,$$

$$\Lambda = \Gamma, \Gamma_1, \Gamma_2, \quad \text{or} \quad \Gamma_3 \cup \Gamma', \quad (4.5)$$

where F_τ is the circulation along any island boundary. It is difficult to use (4.5) if F_τ is unknown.

The constants C_1, C_2, and C_3 are directly calculated from the drifter data together with the spectral coefficients in (4.2). Notice that (4.5) can be easily included into the reconstruction process using Lagrangian multipliers.

f. Regression equations

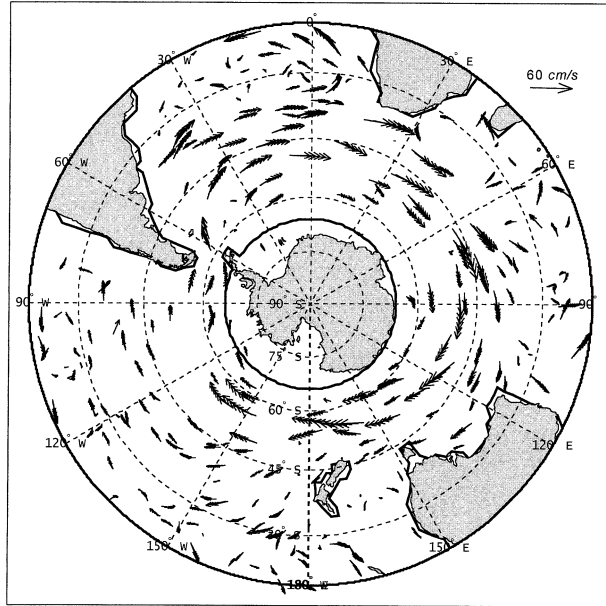For domain A, the linear regression is used in the reconstruction. For domain B, the nonlinear regression

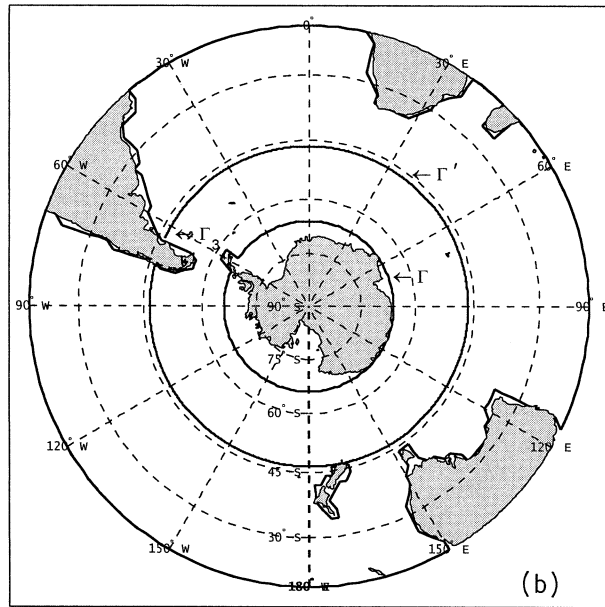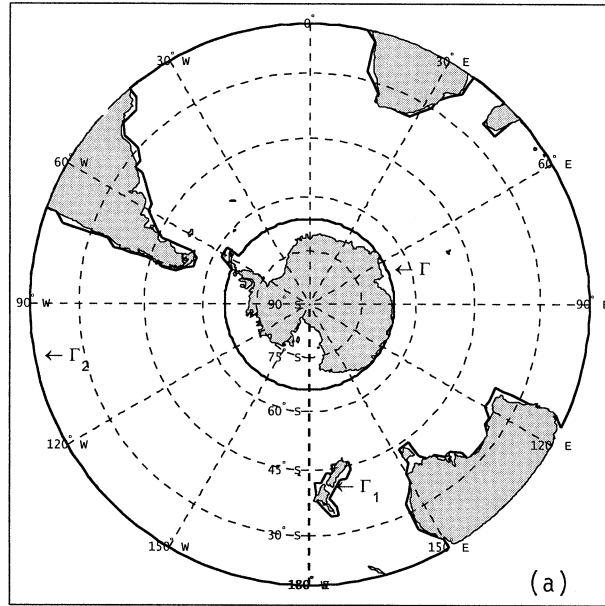FIG. 21. Lagrangian velocities along the smoothed buoy trajectories in May 1979.

FIG. 22. (a) Domain A with two rigid boundaries Γ, Γ_2, and an island Γ_1; and (b) domain B with two rigid boundaries Γ, Γ_3, and one open boundary Γ'.

model [(4.2)–(4.5) of Part I] is used since the basis functions $\{\Psi_k\}$ depend on the constant C_3.

g. Reconstructed circulation in domain A

Domain A is enclosed by rigid boundaries. The basis functions of both potentials are obtained using the rigid boundary conditions [(2.7) of Part I]. For domain A, 1596 observations are analyzed from 213 buoys, and the VCCF is used to determine the model parameters (K_{opt}, M_{opt}) by (30, 4). The reconstructed circulation (Fig. 23) shows the main characteristics of the Southern Ocean circulation, such as the Antarctic Circumpolar Current, the Brazil–Malvinas Confluence, the Agulhas Current, the West Australian Current, as well as the South Pacific gyre, the Indian Ocean gyre, and the South Atlantic gyre. The Antarctic Circumpolar Current meanders and has several branches extending into the Pacific Ocean. The mean current speed over the Southern Ocean is about 20 cm s^{-1}. Since the data are only on the sea surface, it is not estimated the volume transport in the Southern Ocean, such as through the Drake Passage.

h. Reconstructed circulation in domain B

Domain B is a subarea of domain A, and has an open boundary along 48°S. The iteration process [(4.2)–(4.4) of Part I] are used to get the spectral coefficients. The boundary extension method is used to obtain the first guess of the spectral coefficients $\delta\mathbf{h}^{[0]}$ for the iteration. The calculated velocity along the open boundary (48°S) is used to determine the coefficient κ for this domain.

For the reconstruction of domain B, 638 observations from 112 buoys are analyzed; VCCF is used to determine the optimal mode truncation $(K_{opt}, M_{opt}, S_{opt})$ as (30, 4, 7); and a seventh-order polynomial is used to smooth k along the open boundary.

Thus, two velocity fields for domain B are obtained: (a) a subset of the domain A velocity field (Fig. 24a), and (b) a reconstructed field for domain B, only with an open
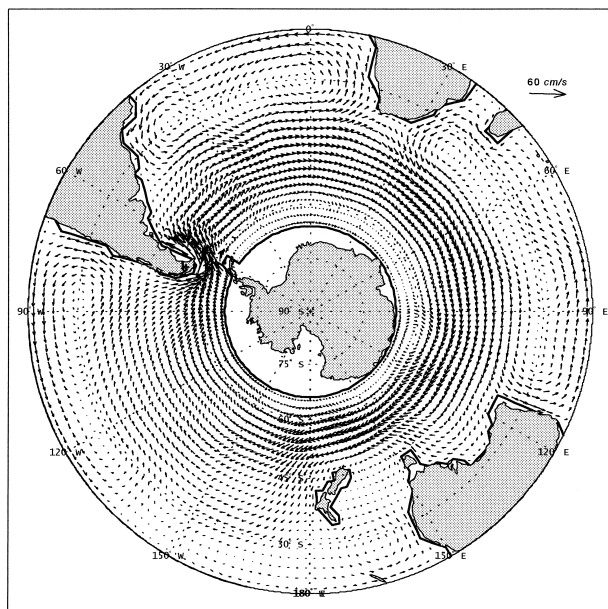
FIG. 23. Reconstructed circulation in domain A.

boundary at 48°S (Fig. 24b). The two flow fields are quite similar. The residue (Fig. 25) between the two fields (subset of domain A minus domain B) is quite weak, and the relative root-mean-square difference between the two fields is calculated using (5.10) of Part I:

$$\chi_u = 0.09 \quad \text{and} \quad \chi_v = 0.13, \qquad (4.7)$$

which indicates that the scheme leads to 9% to 13% of errors in processing the drifter buoy data.

The reconstructed normal and tangential velocities along 48°S (open boundary for domain B) obtained from domain A's reconstruction (denoted by "1") using the rigid boundary conditions (for a larger domain) and domain B's reconstruction (denoted by "2") using the iteration procedure [(4.2)–(4.4) of Part I] are examined (Fig. 26). The difference between the two fields for both components is quite small, except for the tangential component in the South Pacific sector (150°E–100°W). The maximum difference is found at 165°W: the tangential velocity is 25.5 cm s^{-1} from domain A's estimation, and about 13 cm s^{-1} from domain B's estimation.

5. Conclusions

First, the results from three experiments are presented to illustrate the capability of the theoretical approach developed in Part I in flow reconstruction from noisy and sparse data: (a) reconstruction of three-dimensional velocity at 20-m depth in the eastern Black Sea, (b) reconstruction of circulation pattern in a wind driven shallow water basin flow with open boundary, and (c) reconstruction of large-scale circulation in the Southern Ocean from the FGGE drifter data.
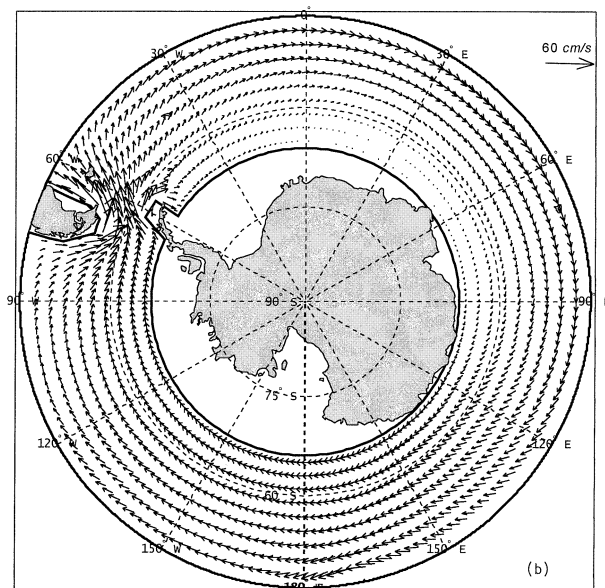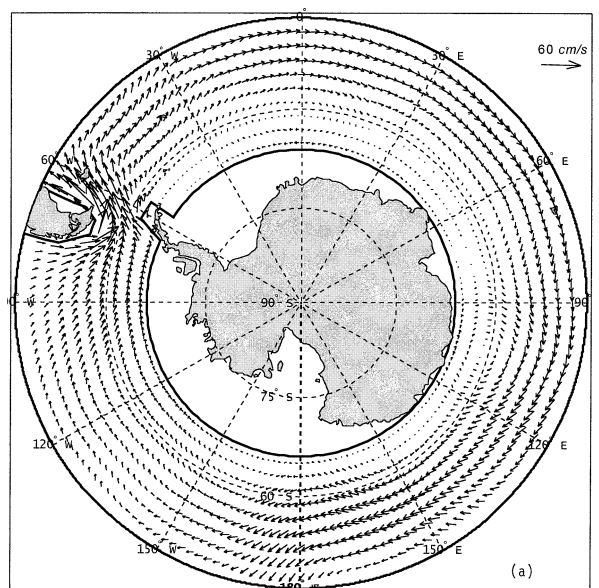

FIG. 24. Reconstructed circulation in domain B (a) obtained from Fig. 23 and (b) calculated using the open boundary condition (4.10) of Part I.

Second, the difficult problem in any reconstruction is how to select the optimal field from noisy data. In oceanography, the low-order noise statistics are usually unknown. This difficulty is overcome using a special regularization algorithm. The three cases presented demonstrate the capability of the algorithm in reconstruction with the noise-to-signal ratio and the condition number up to 1 and 10^8 to 10^{10}, respectively.

Third, the results from the three cases demonstrate that the sampling strategy is the key issue for the reconstruction skill. In the eastern Black Sea experiment, the reconstruction skill deteriorates with the reduction
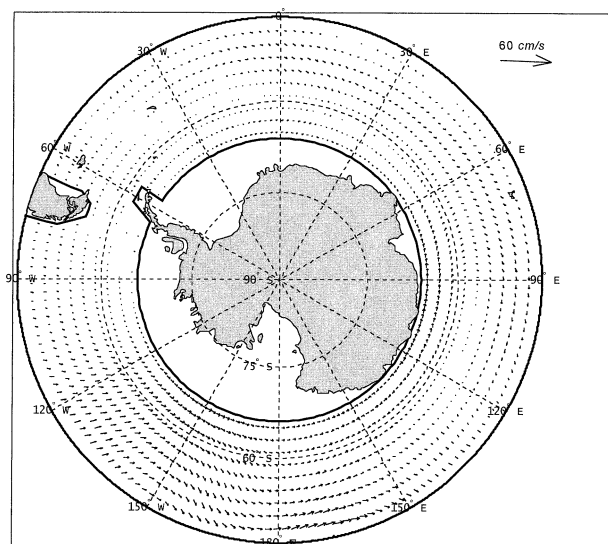
FIG. 25. Residual velocity vectors between the two calculations.

FIG. 26. (a) Normal and (b) tangential velocities along Γ' (48°S) obtained from the calculation for domain A (denoted by the symbol "1"), and from direct calculation for domain B using the open boundary condition (2.10) of Part I (denoted by the symbol "2").

of sampling density from high sampling density (19 observations) to medium sampling density (14 observations), and from medium sampling density to low sampling density (9 observations). For the low sampling density, we can only estimate six toroidal modes and one poloidal mode. Any combination of these modes results to a poor reconstruction skill.

Fourth, the empirical functional J_{emp} is not the best cost function for estimating the optimal mode truncation of the toroidal and poloidal potentials or other free external model parameter from limited data. It is recommended to use the Vapnik–Chervonkis approach (Vapnik 1983) instead of J_{emp} since the VCCF is proven in the three cases to be robust to the variations of the sampling size.

Fifth, reconstruction skill depends on free model parameters ($K_{\text{opt}}, M_{\text{opt}}, S_{\text{opt}}$). The three cases show that the reconstruction of the normal velocity at the open boundary is more accurate than the tangential velocity. However, this result might not be universal since the reconstruction skill of normal and tangential velocities depends on the flow structure along the open boundary.

Sixth, the reconstruction process is reduced to a linear regression model for known and unknown simplified boundary conditions if the calculation is over a simply connected domain. For a multiply connected domain with unknown or known boundary conditions, the flow reconstruction is only possible using a nonlinear regression model with independent parameterization for boundary conditions.

Acknowledgments. The authors are grateful to Professor Christopher N. K. Mooers, at the University of Miami for invaluable suggestions. Ivanov thanks Professor A. D. Kirwan Jr. for fruitful discussions about the problem of open boundary conditions. The authors
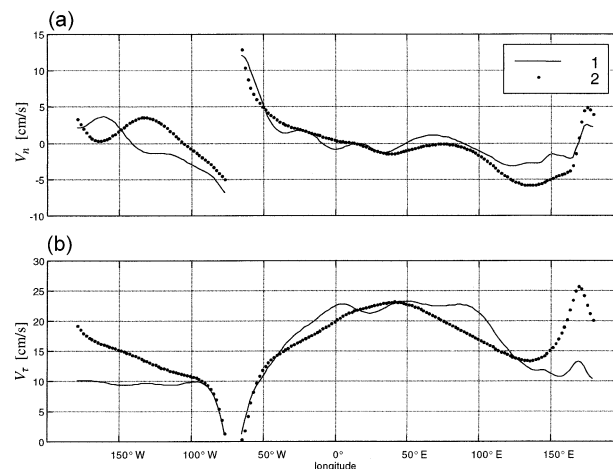
REFERENCES

Blumberg, A. F., and G. L. Mellor, 1987: A description of a three-dimensional coastal ocean circulation model. *Three Dimensional Coastal Ocean Models,* Amer. Geophys. Union, 1–16.

Chu, P. C., 1999: Fundamental circulation functions for determination of open boundary conditions. *Proc. Third Conf. on Coastal Atmospheric and Oceanic Prediction and Processes,* Phoenix, AZ, Amer. Meteor. Soc., 389–394.

——, C. Fan, and L. L. Ehret, 1997: Determination of open boundary conditions with an optimization method. *J. Atmos. Oceanic Technol.,* **14,** 1066–1071.

——, L. M. Ivanov, L. H. Kantha, O. V. Melnichenko, and Y. A. Poberezhny, 2002: Power law decay in model predictability skill. *Geophys. Res. Lett.,* **29** (15), 1748, doi:10.1029/2002GL014891.

——, ——, T. P. Korzhova, T. M. Margolina, and O. V. Melnichenko, 2003: Analysis of sparse and noisy ocean current data using flow decomposition, Part I: Theory. *J. Atmos. Oceanic Technol.,* **20,** 478–491.

Daniault, N., and Y. Menard, 1985: Eddy kinetic energy distribution in the Southern Ocean from altimetry and FGGE drifting buoys. *J. Geophys. Res.,* **90,** 11 877–11 889.

Eremeev, V. N., and L. M. Ivanov, 1987: *Tracers in Ocean: Parameterization and Modeling Dynamics* (in Russian). K. Naukova Dumka, 147 pp.

——, ——, S. V. Kochergin, and O. V. Melnichenko, 1992: Seasonal variability and types of currents in the upper layer of the Black Sea. *Sov. J. Phys. Oceanogr.,* **3,** 193–208.

Eykhoff, P., 1973: *System Identification: Parameter and State Estimation.* Elsevier, 555 pp.

Flierl, G. R., 1977: Simple applications of McWilliams' ''A note on a consistent quasi-geostrophic model in a multiply connected domain.'' *Dyn. Atmos. Ocean, 1,* 443–453.

Hoffmann, E. E., 1985: The large-scale horizontal structure of the Antarctic Circumpolar Current from FGGE drifters. *J. Geophys. Res.,* **90,** 7087–7097.

Garrett, J. C., 1980: Availability of the FGGE drifting buoy system data set. *Deep-Sea Res.,* **27A,** 1083–1086.

Kamenkovich, V. M., 1961: The integration of the marine current theory equations in multiply connected regions (in Russian). *Dokl. Akad. Nauk USSR,* **138,** 629–631.

Knysh, V. V., S. G. Demishev, and G. K. Korotaev, 2001: On model reconstruction of climate seasonal variability of the Black Sea (in Russian). *Mar. Hydrophys. J.,* **3,** 221–227.

Ivanov, L. M., A. D. Kirwan Jr., and T. M. Margolina, 2001: Filtering noise from oceanographic data with some applications for the Kara and Black Seas. *J. Mar. Sys.,* **28,** 113–139.

MATLAB, 1997: *High-performance numeric computation and visualization software,* The Math Works, Inc., 548 pp.

McWilliams, J. C., 1977: A note on a consistent quasi-geostrophic model in a multiply connected domain. *Dyn. Atmos. Ocean,* **1,** 427–441.

Patterson, S. L., 1985: Surface circulation and kinetic energy distributions in the Southern Hemisphere oceans from FGGE drifting buoys. *J. Phys. Oceanogr.,* **15,** 865–884.

Platzman, G. W., 1979: Effects of multiple connectivity on a finite-element barotropic model. *J. Phys. Oceanogr.,* **9,** 1276–1283.

Sabelfeld, K., 1991: *Monte-Carlo Methods in Boundary Value Problems.* Springer-Verlag, 274 pp.

Shulman, I., and Coauthors, 2000: Development of the high resolution, data assimilating numerical model of the Monterey Bay. *Proc. Estuarine and Coastal Modeling,* New Orleans, LA, ASCE, 980–994.

Shultz, W. J., Jr., 1999: Ocean surface maps from blending disparate data through normal mode analysis. Tech. Rep. 99-02, Old Dominion University, 137 pp.

Staneva, J. V., E. V. Stanev, and R. H. Rachev, 1995: Heat balance estimates using atmospheric analysis data: A case of the Black Sea. *J. Geophys. Res.,* **100,** 18 581–18 596.

Tikhonov, A. N., A. V. Goncharsky, V. V. Stepanov, and A. G. Yagola, 1990: *Numerical Methods for Solving Ill-Posed Problems* (in Russian). Nauka, 223 pp.

Vapnik, V. H., 1983: *Reconstruction of Empirical Laws from Observations* (in Russian). Nauka, 447 pp.